1. INTRODUCTION

1.1. PURPOSE AND SCOPE OF THE GUIDELINES

These guidelines revise and replace United States Environmental Protection Agency (EPA) Guidelines for Carcinogen Risk Assessment published in 51 FR 33992, September 24, 1986. The guidelines provide EPA staff and decision makers with guidance for developing and using risk assessments. They also provide basic information to the public about the Agency's risk assessment methods. These guidelines are used with other risk assessment guidelines that the Agency has developed, such as the Mutagenicity Risk Assessment Guidelines (U.S. EPA, 1986c). and the Exposure Assessment Guidelines (U.S. EPA, 1992a). Consideration of other Agency guidance documents is particularly important when procedures for evaluating specific target organ effects have been developed (e.g., assessment of thyroid follicular cell tumors (U.S. EPA, 1998a)), or when there is a concern for a particular sensitive subpopulation for which the Agency has developed guidance, for example, EPA Guidelines for Developmental Toxicity Risk Assessment (U.S. EPA, 1991d). These guidelines discuss hazards to children that may result from exposures during preconception, prenatal, or postnatal development to sexual maturity. Similar guidelines exist for Reproductive Toxicant Risk Assessment (U.S. EPA, 1996c) and for Neurotoxicity Risk Assessment (U.S. EPA, 1998c). All of these guidelines should be consulted when conducting a risk assessment in order to insure that information from studies on carcinogenesis and other health effects are considered together in the overall characterization of risk. This is particularly true in the case in which a precursor effect to tumor is also a precursor or endpoint of other health effects and is used in dose-response assessment. The overall characterization of risk will be the basis for carrying out assessments of instances in which fetuses, infants, or children are at risk or disproportionately affected by economically significant Agency actions. Characterization for the latter purpose is outlined in the Agency guidance by the Office of Children's Health Protection to carry out E.O. 13045, "Protection of Children From Environmental Health Risks and Safety Risks" issued on April 21, 1997.

The guidelines encourage both regularity in procedures to support consistency in scientific components of Agency decision making and innovation to remain up-to-date in scientific thinking. In balancing these goals, the Agency relies on established scientific peer review processes (EPA, 1998b). The guidelines incorporate basic principles and science policies based on evaluation of the currently available information. As more is discovered about carcinogenesis, the need will arise to make appropriate changes in risk assessment guidance. The Agency will revise these guidelines when extensive changes are due. In the interim, the Agency will issue special reports,

after appropriate peer review, to supplement and update guidance on single topics, (e.g., U.S. EPA, 1991b). The incorporation of new, peer-reviewed scientific understanding and data in an assessment is always consistent with the purposes of these guidelines.

1.2. ORGANIZATION AND APPLICATION OF THE GUIDELINES

1.2.1. Organization

Publications of the Office of Science and Technology Policy (OSTP, 1985) and the National Research Council (NRC, 1983, 1994) provide information and general principles about risk assessment. Risk assessment uses available scientific information on the properties of an agent[1] and its effects in biological systems to provide an evaluation of the potential for harm as a consequence of environmental exposure. The 1983 and 1994 NRC documents organize risk assessment information into four areas: hazard identification, dose-response assessment, exposure assessment, and risk characterization. This structure appears in these guidelines, which additionally emphasize characterization of evidence and conclusions in each part of the assessment. In particular, the guidelines adopt the approach of the NRC's 1994 report in adding a dimension of characterization to the hazard identification step. Added to the identification of hazard is an evaluation of the conditions under which its expression is anticipated. The risk assessment questions addressed in these guidelines are:

- For hazard--Can the agent present a carcinogenic hazard to humans, and if so, under what circumstances?
- For dose-response--At what levels of exposure might effects occur?
- For exposure--What are the conditions of human exposure?
- For risk--What is the character of the risk? How well do data support conclusions about the nature and extent of the risk?

1.2.2. Application

The guidelines apply within the framework of policies provided by applicable EPA statutes and do not alter such policies. The guidelines cover assessment of available data. They do not imply that one kind of data or another is prerequisite for regulatory action concerning any agent. Risk management applies directives of regulatory legislation, which may require consideration of

---

[1]The term "agent" refers generally to any chemical substance, mixture, or physical or biological entity being assessed, unless otherwise noted (See sec. 1.2.2 for a note on radiation.).

potential risk, or solely hazard or exposure potential, along with social, economic, technical, and other factors in decision making. Risk assessments support decisions, but to maintain their integrity as decision making tools, they are not influenced by consideration of the social or economic consequences of regulatory action.

The assessment of risk from radiation sources is based on continuing examination of human data by the National Academy of Sciences/National Research Council in its series of numbered reports: “Biological Effects of Ionizing Radiation”. While the general principles of these guidelines apply to radiation risk assessments, their details are most focused on other kinds of agents. They do not attempt to guide the ongoing conduct of radiation risk assessment.

Not every EPA assessment has the same scope or depth. Agency staff often conduct screening-level assessments for priority-setting or separate assessments of hazard or exposure for ranking purposes or to decide whether to invest resources in collecting data for a full assessment. Moreover, a given assessment of hazard and dose-response may be used with more than one exposure assessment that may be conducted separately and at different times as the need arises in studying environmental problems in various media. The guidelines apply to these various situations in appropriate detail given the scope and depth of the particular assessment. For example, a screening assessment may be based almost entirely on structure-activity relationships and default assumptions. As more data become available, assessments can replace or modify default assumptions accordingly. These guidelines do not require that all of the kinds of data covered here be available for either assessment or decision making. The level of detail of an assessment is a matter of Agency management discretion regarding applicable decision making needs.

1.3. USE OF DEFAULT ASSUMPTIONS

The National Research Council, in its 1983 report on the science of risk assessment (NRC, 1983), recognized that default assumptions are necessarily made in risk assessments where gaps exist in general knowledge or in available data for a particular agent. These default assumptions are inferences based on general scientific knowledge of the phenomena in question and are also matters of policy concerning the appropriate way to bridge uncertainties that concern potential risk to human health (or, more generally, to environmental systems) from the agent under assessment.

EPA's 1986 guidelines for cancer risk assessment (EPA, 1986b) were developed to be responsive to the principles of the 1983 NRC report. The guidelines contained a number of default assumptions. They also encouraged research and analysis that would lead to new risk

assessment methods and data and anticipated that these would replace defaults. The 1986 guidelines did not explicitly discuss how to depart from defaults.

In its 1994 report on risk assessment, the NRC supported continued use of default assumptions (NRC, 1994). The NRC report thus validated a central premise of the approach to risk assessment that EPA had evolved in preceding years--the making of science policy inferences to bridge gaps in knowledge--while at the same time recommending that EPA develop more systematic and transparent guidelines to inform the public of the default inferences EPA uses in practice. It recommended that the EPA review and update the 1986 guidelines in light of evolving scientific information and experience in practice in applying those guidelines, and that the EPA explain the science and policy considerations underlying current views as to the appropriate defaults and provide general criteria to guide preparers and reviewers of risk assessments in deciding when to depart from a default.

1.3.1. Default Assumptions

The 1994 NRC report contains several recommendations regarding flexibility and the use of default options:

- EPA should continue to regard the use of default options as a reasonable way to deal with uncertainty about underlying mechanisms in selecting methods and models for use in risk assessment.
- EPA should explicitly identify each use of a default option in risk assessments.
- EPA should clearly state the scientific and policy basis for each default option.
- The Agency should consider attempting to give greater formality to its criteria for a departure from default options in order to give greater guidance to the public and to lessen the possibility of ad hoc, undocumented departures from default options that would undercut the scientific credibility of the Agency's risk assessments. At the same time, the Agency should be aware of the undesirability of having its guidelines evolve into inflexible rules.
- EPA should continue to use the Science Advisory Board and other expert bodies. In particular, the Agency should continue to make the greatest possible use of peer review, workshops, and other devices to ensure broad peer and scientific participation to guarantee that its risk assessment decisions will be based on the best science available through a process that allows full public discussion and peer participation by the scientific community.

In the 1983 report (p. 28), NAS defined the use of "inference options" (default options) as

a means to bridge inherent uncertainties in risk assessment. These options exist when the assessment encounters either "missing or ambiguous information on a particular substance" or "gaps in current scientific theory." Since there is no instance in which a set of data on an agent or exposure is complete, all risk assessments must use general knowledge and policy guidance to bridge data gaps. Animal toxicity data are used, for example, to substitute for human data because we do not test human beings. The report described the components of risk assessment in terms of questions encountered during analysis for which inferences must be made. The report noted (p. 36) that many components ". . . lack definitive scientific answers, that the degree of scientific consensus concerning the best answer varies (some are more controversial than others), and that the inference options available for each component differ in their degree of conservatism. The choices encountered in risk assessment rest, to various degrees, on a mixture of scientific fact and consensus, on informed scientific judgment, and on policy determinations (the appropriate degree of conservatism). . . ." The report did not note that the mix varies significantly from case to case. For instance, a question that arises in hazard identification is how to use experimental animal data when the routes of exposure differ between animals and humans. A spectrum of inferences could be made: The most protective, or risk adverse one is that effects in animals from one route may be seen in humans by another route. An intermediate one is a conditional inference that such translation of effects will be assumed if the agent is absorbed by humans through the second route. A nonprotective one that no inference is possible and the agent's effects in animals must be tested by the second route. The choice of an inference, as the report observed, comes from more than scientific thinking alone. While the report focused mainly on the idea of conservatism of public health as a science policy rationale for making the choice, it did not evaluate other considerations.

These revised guidelines retain the use of default assumptions as recommended in the 1994 report. Since the primary goal of EPA actions is public health protection and that, accordingly, as an Agency policy, the defaults used in the absence of scientific data to the contrary have been chosen to be health protective. The defaults described below remain public health conservative when applied in combination in risk assessment, however, any individual default may not constitute the most conservative position vis-a-vis that position. To do so would lead to risk assessments that far exceed the actual risks and this would not be in keeping with the principles discussed in the NAS 1994 report.

In addition, the guidelines reflect evaluation of experience in practice in applying defaults and departing from them in individual risk assessments conducted under the 1986 guidelines. The application and departure from defaults and the principles to be used in these judgments have been

matters of debate among practitioners and reviewers of risk assessments. The guidelines here are intended to be both explicit and more flexible than in the past concerning the basis for making departures from defaults, recognizing that expert judgment and peer review are essential elements of the process.

In response to the recommendations of the 1994 report, these guidelines call for identification of the default assumptions used within assessments and for highlighting significant issues about defaults within characterization summaries of component analyses in assessment documents. As to the use of peer review to aid in making judgments about applying or departing from defaults, we agree with the NRC recommendation. The Agency has long made use of workshops, peer review of documents and guidelines, and consultations as well as formal peer review by the Science Advisory Board (SAB). In 1998, the Administrator of EPA published a peer review guidance for EPA scientific work products that increases the amount of peer review for risk assessments as well as other work, continuing a series of guidance actions in response to the NRC report and to SAB recommendations (U.S. EPA, 1994b, 1997b, 1998b).

The 1994 NRC report recommended that EPA should consider adopting principles or criteria that would give greater formality and transparency to decisions to depart from defaults. The report named several possible criteria for such principles (p. 7): ". . . [P]rotecting the public health, ensuring scientific validity, minimizing serious errors in estimating risks, maximizing incentives for research, creating an orderly and predictable process, and fostering openness and trustworthiness. There might be additional relevant criteria. . . ." The report indicated, however, that the committee members had not reached consensus on a single criterion to address the key issue of how much certainty or proof a risk assessor must have in order to justify departing from a default. Appendix N of the report contains two presentations of alternative views held by some committee members on this issue. One view, known as "plausible conservatism," suggested that departures from defaults should not be made unless new information improves the understanding of a biological process to the point that relevant experts reach consensus that the protective default assumption concerning that process is no longer plausible. The same criterion was recommended where the underlying scientific mechanism is well understood, but where a default is used to address missing data. In this case, the default should not be replaced with case-specific data unless it is the consensus of relevant experts that the proffered data make the default assumption no longer plausible. Another view, known as the "maximum use of scientific information" approach, acknowledged that the initial choice of defaults should be protective, but argued that conservatism should not be a factor in determining whether to depart from the default in favor of an alternate biological theory or alternate data. According to this view, it should not

be necessary to reach expert consensus that the default assumption had been rendered implausible; it should be sufficient that risk assessors find the alternate approach more plausible than the default.

The EPA is not adopting a general list of formal decision criteria in the sense of a checklist applicable to departures from defaults. It would not be helpful to generate a checklist of uniform criteria. Risk assessments are highly variable in content and purpose. Screening assessments may be purposely "worst case" in their default assumptions to eliminate problems from further investigation. Subsequent risk assessments based on a fuller data set can discard worst-case default assumptions in favor of plausibly protective assumptions and progressively replace or modify the latter with data. No uniform checklist will fit all cases or all kinds of data. Moreover, some departures from defaults are controversial, some are not. Generic checklists would likely become more a source of rote discussion than of enlightenment about the process.

Nonetheless, for one issue, the EPA has adopted principles to give greater formality and transparency to decisions to depart from defaults. The EPA has developed a framework for evaluating a postulated mode of action which appears in section 2.5, below. The use of mode of action information to make decisions about human relevance of animal data, to assist in identifying sensitive subpopulations, and to decide upon approaches to high dose to low dose extrapolation in dose-response assessment is a fundamental part of these guidelines. The framework of section 2.5. contains principles derived from Bradford Hill criteria for considering causation in human epidemiologic studies and is meant to weigh the question whether empirical data support a mode of action that is proposed in a particular case.

The guidelines use a combination of principles and process in the application of and departure from default assumptions. The framework of default assumptions allows risk assessment to proceed when current scientific theory or available case-specific data do not provide firm answers in a particular case, as the 1983 NRC report outlined. Some of the default assumptions bridge large gaps in fundamental knowledge which will be filled by basic research on the causes of cancer and on other biological processes, rather than by agent-specific testing. Other default assumptions bridge smaller data gaps that can feasiblely be filled for a single agent, such as whether a metabolic pathway in test animals is like (default) or unlike that in humans.

The decision to use a default, or not, is a choice considering available information on an underlying scientific process and agent-specific data, depending on which kind of default it is. Generally, if a gap in basic understanding exists, or if agent-specific data are missing, the default is used without pause. If data are present, their evaluation may reveal inadequacies that also lead to use of the default. If data support a plausible alternative to the default, but no more strongly than

they support the default, both the default and its alternative are carried through the assessment and characterized for the risk manager. If the alternative to the default are strongly supported by data, the alternative may be used in place of the default. These guidelines provide a framework for making such decisions. Note that, as discussed above, there is a spectrum of difficulty in replacing default positions with empirical data. In the case of showing a mode of action, there is need for extensive experimentation to support an hypothesis as to mode of action for a specific tumor response, including coverage of the issue whether other modes of action are plausible.

Note that screening assessments may appropriately use "worst case" inferences to determine if, even under those conditions, risk is low enough that a problem can be eliminated from further consideration.

Scientific peer review, peer consultative workshops and similar processes are the principal ways determining the strength of thinking and generally accepted views within the scientific community about the application of and departure from defaults and about judgments concerning the plausibility and persuasiveness of data in a particular case.

The discussion of major defaults below together with the explicit discussion of the choice of inferences within the assessment and the processes of peer review and peer consultation (U.S. EPA, 1998b) will serve the several goals stated in the 1994 NRC report. One is to encourage research, since results of research efforts will be considered. Another is to allow timely decision making, when time is a constraint, by supporting completion of the risk assessment using defaults as needed. Another is to be flexible, using new science as it develops. Finally, the use of public processes of peer consultation and peer review will ensure that discipline of thought is maintained to support trust in assessment results.

There is no one set of rules for making the judgment of whether a data analysis is both biologically plausible and persuasive as applied to the case at hand. Two criteria that apply in these guidelines are that the underlying scientific principle has been generally accepted within the scientific community and that supportive experiments are available that test the application of the principle to the agent under review. For example, mutagenicity through reactivity with DNA has been generally accepted as a carcinogenic influence for many years. This acceptance, together with evidence of such mutagenicity in experiments on an agent, provides plausible and persuasive support for the inference that mutagenicity is a mode of action for the agent.

Judgments about plausibility and persuasiveness of analyses vary according to the scientific nature of the default. An analysis of data may replace a default or modify it. An illustration of the former is development of EPA science policy on the issue of the relevance for humans of male rat kidney neoplasia involving alpha 2u globulin (U.S. EPA, 1991b). The 1991

EPA policy gives guidance on the kind of experimental findings that demonstrate whether the alpha 2u globulin mechanism is present and responsible for carcinogenicity in a particular case. Before this policy guidance was issued, the default assumption was that neoplasia in question was relevant to humans and indicated the potential for hazard to humans. A substantial body of data was developed by public and private research groups as a foundation for the view that the alpha 2u globulin-induced response was not relevant to humans. These studies first addressed the alpha 2u globulin mechanism in the rat and whether this mechanism has a counterpart in the human being, both were large research efforts. The resulting data presented difficulties; some reviewers were concerned that the mechanism in the rat appeared to be understood only in outline, not in detail, and felt that the data were insufficient to show the lack of a counterpart mechanism in humans. It was particularly difficult to support a negative such as the nonexistence of a mechanism in humans because so little is known about what the mechanisms are in humans. Despite these concerns, in its 1991 policy guidance, EPA concluded that the alpha 2u globulin-induced response in rats should be regarded as not relevant to humans (i.e., as not indicating human hazard).

One conclusion from the development and peer review of this policy is that if the default concerns an inherently complex biological question such as mode of action, large amounts of work will be required to replace the default. A second is that "proof" in the strict sense of having proved a negative is neither reasonable nor required. Rather the alternative may displace the default when it is supported by clear and convincing evidence and is generally accepted in peer review. The issue of relevance may not always be so difficult. It would be an experimentally easier task, for example, to determine whether carcinogenesis in an animal species is due to a metabolite of the agent in question that is not produced in humans.

When scientific processes are understood but case-specific data are missing, defaults can be constructed to be modified by experimental data, even if data do not suffice to replace them entirely. For example, the approaches adopted in these guidelines for scaling dose from experimental animals to humans are constructed to be either modified or replaced as data become available on toxicokinetic parameters for the particular agent being assessed. Similarly, the selection of an approach or approaches for dose-response assessment is based on a series of decisions that consider the nature and adequacy of available data in choosing among alternative modeling and default approaches.

The 1994 NRC report notes (p. 6) that "[a]s scientific knowledge increases, the science policy choices made by the Agency and Congress should have less impact on regulatory decision making. Better data and increased understanding of biological mechanisms should enable risk

assessments that are less dependent on protective default assumptions and more accurate as predictions of human risk." Undoubtedly, this is the trend as scientific understanding increases. However, some gaps in knowledge and data will doubtless continue to be encountered in assessment of even data-rich cases, and it will remain necessary for risk assessments to continue using defaults within the framework set forth here.

1.3.2. Major Defaults

This discussion covers the major default assumptions commonly employed in a cancer risk assessment and adopted in these guidelines. They are predominantly inferences necessary to use data observed under empirical conditions to estimate events and outcomes under environmental conditions. Several inferential issues arise when effects seen in a subpopulation of humans or animals are used to infer potential effects in the population of environmentally exposed humans. Several more inferential issues arise in extrapolating the exposure-effect relationship observed empirically to lower-exposure environmental conditions. The following issues cover the major default areas. Typically, an issue has some sub-issues; they are introduced here, but are discussed in greater detail in later sections.

- Is the presence or absence of effects observed in a human population predictive of effects in another exposed human population?
- Is the presence or absence of effects observed in an animal population predictive of effects in exposed humans?
- How do metabolic pathways relate across species? Among different age groups, between sexes in humans?
- How do toxicokinetic processes relate across species? Among different age groups, between sexes in humans?
- What is the correlation of the observed dose-response relationship to the relationship at lower doses?

1.3.2.1. *Is the Presence or Absence of Effects Observed in a Human Population Predictive of Effects in Another Exposed Human Population?*

When cancer effects in exposed humans are attributed to exposure to an exogenous agent, the default assumption is that such data are predictive of cancer in any other exposed human population. Studies either attributing cancer effects in humans to exogenous agents or reporting no effects are often studies of occupationally exposed humans. By sex, age, and general health, workers are not representative of the general population exposed environmentally to the

same agents. In such studies there is no opportunity to observe those who are likely to be under represented, e.g., fetuses, infants and children, women, or people in poor health, who may respond differently from healthy workers. Therefore, it is understood that this assumption could still underestimate the response of certain human subpopulations. (NRC, 1993a, 1994).

There is not enough knowledge yet to form a basis for any generally applicable, qualitative or quantitative inference to compensate for this knowledge gap. In these guidelines, this problem is left to analysis in individual cases, to be attended to with further general guidance as future research and information allow. When information on a sensitive subpopulation exists, it will be used. For example, an agent such as diethylstilbestrol (DES)causes a rare form of vaginal cancer (clear-cell adenocarcinoma) (Herbst, 1971) in about 1 per thousand of adult women whose mothers were exposed during pregnancy (Hatch et al., 1998). *When cancer effects are not found in an exposed human population, this information by itself is not generally sufficient to conclude that the agent poses no carcinogenic hazard to this or other populations of potentially exposed humans including sensitive subpopulations.* This is because epidemiologic studies usually have low power to detect and attribute responses, and typically evaluate cancer potential in a restricted population (e.g., by age, occupation, etc.). The topic of susceptibility and variability is addressed further in the discussion of quantitative default assumptions about dose-response relationships below.

1.3.2.2. *Is the Presence or Absence of Effects Observed in an Animal Population Predictive of Effects in Exposed Humans?*

The default assumption is that positive effects in animal cancer studies indicate that the agent under study can have carcinogenic potential in humans. Thus, if no adequate human data are present, positive effects in animal cancer studies are a basis for assessing the carcinogenic hazard to humans. This assumption is a public health conservative policy, and it is both appropriate and necessary given that we do not test for carcinogenicity in humans. The assumption is supported by the fact that nearly all of the agents known to cause cancer in humans are carcinogenic in animals in tests with adequate protocols (IARC, 1994; Tomatis et al., 1989; Huff, 1994). Moreover, almost one-third of human carcinogens were identified subsequent to animal testing (Huff, 1993). Further support is provided by research on the molecular biology of cancer processes, which has shown that the mechanisms of control of cell growth and differentiation are remarkably homologous among species and highly conserved in evolution. Nevertheless, the same research tools that have enabled recognition of the nature and commonality of cancer processes at the molecular level also have the power to reveal differences

and instances in which animal responses are not relevant to humans (Linjinsky, 1993; U.S. EPA,1991b). Under these guidelines, available mode of action[2] information is studied for its implications in both hazard and dose-response assessment and its effect on default assumptions.

There may be instances in which the use of an animal model would identify a hazard in animals that is not truly a hazard in humans (e.g., the alpha-2u-globulin association with renal neoplasia in male rats (U.S. EPA, 1991b)). The extent to which animal studies may yield false positive indications for humans is a matter of scientific debate. To demonstrate that a response in animals is not relevant to any human situation, adequate data to assess the relevancy issue must be available.

The default assumption is that effects seen at the highest dose tested are appropriate for assessment, but it is necessary that the experimental conditions be scrutinized. Animal studies are conducted at high doses in order to provide statistical power, the highest dose being one that is minimally toxic (maximum tolerated dose). Consequently, the question often arises whether a carcinogenic effect at the highest dose may be a consequence of cell killing with compensatory cell replication or of general physiological disruption, rather than inherent carcinogenicity of the tested agent. There is little doubt that this may happen in some cases, but skepticism exists among some scientists that it is a pervasive problem (Ames and Gold, 1990; Melnick et al., 1993a; Melnick et al., 1993b; Barrett, 1993). If adequate data demonstrate that the effects are solely the result of excessive toxicity rather than carcinogenicity of the tested agent per se, then the effects may be regarded as not appropriate to include in assessment of the potential for human carcinogenicity of the agent. This is a matter of expert judgment, considering all of the data available about the agent including effects in other toxicity studies, structure-activity relationships, and effects on growth control and differentiation.

When cancer effects are not found in well-conducted animal cancer studies in two or more appropriate species and other information does not support the carcinogenic potential of the agent, these data provide a basis for concluding that the agent is not likely to possess human carcinogenic potential, in the absence of human data to the contrary. This default assumption about lack of cancer effects has limitations. It is recognized that animal studies (and

---

[2]Understanding an agent's "mode of action" means understanding the general sequence of events by which it causes effects on cell growth control that result in cancer. "Mode of action" is used rather than "mechanism of action" which is a term that implies complete knowledge of the steps of carcinogenesis at the molecular level, a level of understanding that currently does not exist for any agent.

epidemiologic studies as well) have very low power to detect cancer effects. Detection of a 10% tumor incidence is generally the limit of power with standard protocols for animal studies (with the exception of rare tumors that are virtually markers for a particular agent, e.g., angiosarcoma caused by vinyl chloride). In some situations, the tested animal species may not be predictive of effects in humans; for example, arsenic shows only minimal or no effect in animals, while it is clearly positive in humans. Therefore, it is important to consider other information as well; absence of mutagenic activity or absence of carcinogenic activity among structural analogues, can increase the confidence that negative results in animal studies indicate a lack of human hazard. Another limitation is that standard animal study protocols are not yet available for effectively studying perinatal effects. The potential for effects on the very young generally must be considered separately. Perinatal studies accomplished by modification of existing adult bioassay protocols need to be required in special circumstances under existing Agency policy (U.S. EPA, 1997a,b)

The default assumption is that target organ concordance is not a prerequisite for evaluating the implications of animal study results for humans. Target organs of carcinogenesis for agents that cause cancer in both animals and humans are most often concordant at one or more sites (Tomatis et al., 1989; Huff, 1994). However, concordance by site is not uniform. The mechanisms of control of cell growth and differentiation are concordant among species, but there are marked differences among species in the way control is managed in various tissues. For example, in humans, mutations of the tumor suppressor genes p53 and retinoblastoma are frequently observed genetic changes in tumors. These tumor suppressor genes are also observed to be operating in some rodent tissues, but other growth control mechanisms predominate in other rodent tissues. Thus, an animal response may be due to changes in a control that are relevant to humans, but appear in animals in a different way. However, it is appropriate under these guidelines to consider the influences of route of exposure, metabolism, and, particularly, some modes of action that may either support or not support target organ concordance between animals and humans. When data allow, these influences are considered in deciding whether the default remains appropriate in individual instances (NRC, 1994, p. 121). Another exception to the basic default of not assuming site concordance exists in the context of toxicokinetic modeling. Site concordance is inherently assumed when these models are used to estimate delivered dose in humans based on animal data.

The default is to include benign tumors observed in animal studies in the assessment of animal tumor incidence if they have the capacity to progress to the malignancies with which they are associated. This default is consistent with the approach of the National Toxicology Program

and the International Agency for Research on Cancer and is somewhat more protective of public health than not including benign tumors in the assessment. This treats the benign and malignant tumors as representative of related responses to the test agent (McConnell et al., 1986), which is scientifically appropriate. Nonetheless, in assessing findings from animal studies, a greater proportion of malignancy is weighed more heavily than a response with a greater proportion of benign tumors. Greater frequency of malignancy of a particular tumor type in comparison with other tumor responses observed in an animal study is also a factor to be considered in selecting the response to be used in dose-response assessment.

Benign tumors that are not observed to progress to malignancy are assessed on a case-by-case basis. There is a range of possibilities for their overall significance. They may deserve attention because they are serious health problems even though they are not malignant; for instance, benign tumors may be a health risk because of their effect on the function of a target tissue such as the brain. They may be significant indicators of the need for further testing of an agent if they are observed in a short term test protocol, or such an observation may add to the overall weight of evidence if the same agent causes malignancies in a long term study. Knowledge of the mode of action associated with a benign tumor response may aid in the interpretation of other tumor responses associated with the same agent.

1.3.2.3. *How Do Metabolic Pathways Relate Across Species? Among different age groups, between sexes in humans?*

The default assumption is that there is a similarity of the basic pathways of metabolism and the occurrence of metabolites in tissues in regard to the species-to-species extrapolation of cancer hazard and risk. If comparative metabolism studies were to show no similarity between the tested species and humans and a metabolite(s) were the active form, there would be less support for an inference that the animal response(s) relates to humans. In other cases, parameters of metabolism may vary quantitatively between species; this becomes part of deciding on an appropriate human equivalent dose based on animal studies, optimally in the context of a toxicokinetic model. While the basic pathways are assumed to be the same among humans, the presence of polymorphisms and the maturation of the pathways in infants needs to be considered. The active form of an agent may be present to differing degrees, or completely absent, which may result in greater or lesser risk for subpopulations.

1.3.2.4. *How Do Toxicokinetic Processes Relate Across Species? Among different age groups, between sexes in humans?*

A major issue is how to estimate human equivalent doses in extrapolating from animal studies. *As a default for oral exposure, a human equivalent dose for adults is estimated from data on another species by an adjustment of animal applied oral dose by a scaling factor of body weight to the 0.75 power.* This adjustment factor is used because it represents scaling of metabolic rate across animals of different size. Because the factor adjusts for a parameter that can be improved on and brought into more sophisticated toxicokinetic modeling, when such data become available, the default assumption of 0.75 power can be refined or replaced. *The same factor is used for children because it is slightly more protective than using children's body weight (see section 1.3.5.2).*

For inhalation exposure, a human equivalent dose for adults is estimated by default methodologies that provide estimates of lung deposition and of internal dose. The methodologies can be refined to more sophisticated forms with data on toxicokinetic and metabolic parameters of the specific agent. This default assumption, like the one with oral exposure, is selected in part because it lays a foundation for incorporating better data. *Because of the differences for infants and children, for gases and aerosols, an adjustment is made for their breathing rate and their body weight.* For inhaled particles, the adjustment does not take into account the different size and spacing of airways of children and adults; this difference could result in children and adults retaining particles with a different size distribution and different toxicologic properties. To reduce this uncertainty, EPA is developing a default dosimetry model for children that is based on children's inhalation parameters. The use of information to improve dose estimation from applied to internal to delivered dose is encouraged, including use of toxicokinetic modeling instead of any default, where data are available.

The processes of absorption, distribution, and elimination have important differences among infants, adults, and older adults, e.g., infants tend to absorb metals through the gut more rapidly and more efficiently than older children or adults (Calabrese, 1986). Renal elimination is also not as efficient in infants. While these processes reach adult compentency at about the time of weaning, they may have important implications, particularly when the dose-response relationship for an agent is considered to be nonlinear and there is an exposure scenario disproportionately affecting infants, because in these cases the magnitude of dose is more pertinent than the usual approach in linear extrpolation, of averaging dose across a lifetime. Efficiency of intestinal absorption in older adults tends to be generally less overall for most chemicals. Another notable difference is that, post-weaning (about one year), children have a

higher metabolic rate than adults (Renwick, 1999) and may toxify or detoxify agents at a correspondingly higher rate..

For a route-to-route of exposure extrapolation, *the default assumption is that an agent that causes internal tumors by one route of exposure will be carcinogenic by another route if it is absorbed by the second route to give an internal dose.* This is a qualitative assumption and is considered to be public health conservative. The rationale is that for internal tumors an internal dose is significant no matter what the route of exposure. Additionally, the metabolism of the agent will be qualitatively the same for an internal dose. The issue of quantitative extrapolation of the dose-response relationship from one route to another is addressed case by case. Quantitative extrapolation is complicated by considerations such as first-pass metabolism, but is approachable with empirical data. Adequate data are necessary to demonstrate that an agent will act differently by one route versus another route of exposure.

1.3.2.5. *What Is the Correlation of the Observed Dose-Response Relationship to the Relationship at Lower Doses?*

(To be revised after consideration of comments from SAB January 1999 meeting)

If sufficient data are available, a biologically based model for both the observed range and extrapolation below that range may be used. While no standard biologically based models are in existence, one may be developed if extensive data exist in a particular case and the purpose of the assessment justifies the investment of resources needed. *The default procedure for the observed range of data, when a biologically based model is not used, is to use a curve-fitting model for incidence data.*

In the absence of data supporting a biologically based model for extrapolation outside of the observed range, the choice of approach is based on the view of mode of action of the agent arrived at in the hazard assessment.

The basic default is to assume linearity and use a linear default approach when the mode of action information is supportive of linearity or mode of action is not understood. The linear approach is used when a view of the mode of action indicates a linear response, for example, when a conclusion is made that an agent directly causes alterations in DNA, a kind of interaction that not only theoretically requires one reaction, but also is likely to be additive to ongoing, spontaneous gene mutation. Other kinds of activity may have linear implications, e.g., linear rate-limiting steps, that support a linear procedure also. The linear approach is to draw a straight line between a point of departure from observed data, generally, as a default, the LED_{10}, and the

origin (zero incremental dose, zero incremental response). Other points of departure may be more appropriate for certain data sets; these may be used instead of the LED_{10}. This approach is generally considered to be public health protective. The LED_{10} is the lower 95% limit on a dose that is estimated to cause a 10% response. This level is chosen to account (protectively) for experimental variability. Additionally, it is chosen because it rewards experiments with better designs in regard to number of doses and dose spacing, since these generally will have narrower confidence limits. It is also an appropriate representative of the lower end of the observed range because the limit of detection of studies of tumor effect is about 10%.

The linear default is thought to generally provide an upper bound calculation of potential risk at low doses. e.g., a 1/100,000 to 1/1,000,000 risk; the straight line approach gives numerical results about the same as a linearized multistage procedure. This upper bound is thought to be public health conservative at low doses for the range of human variability considering the typical Agency target range for risk management of 1/1,000,000 to 1/10,000, although it may not completely do so (Bois et al., 1995) if pre-existing disease or genetic constitution place a percentage of the population at risk from any exposure above zero to xenobiotics, natural or manmade. The question of what may be the actual variability in human sensitivity is one that the 1994 NRC report discussed as did the 1993 NRC report on pesticides in children and infants. The NRC has recommended research on the question, and the EPA and other agencies are conducting such research. Given the current state of knowledge, the EPA will assume that the linear default procedure adequately accounts for human variability unless there is case-specific information for a given agent that indicates a particularly sensitive subpopulation, in which case the special information will be used.

When adequate data on mode of action show that linearity is not plausible, and provide sufficient evidence to support a nonlinear mode of action for the general population and any subpopulations of concern, the default changes to a different approach-- a margin of exposure analysis--which assumes that nonlinearity is more reasonable. The departure point is again generally the LED_{10} when incidence data are modeled. When the data available are continuous data such as blood levels of hormones or organ weight, a NOAEL/LOAEL procedure is typically used since modeling approaches for deriving a point of departure from continuous data are not yet available. Until these modeling approaches are developed and adopted, continuous data and data sets that are a mixture of incidence and continuous data can be examined by the NOAEL/LOAEL procedure. In the nonlinear approach, the margin that exists between a human exposure of interest and the point of departure is examined for adequacy to protect public health. A margin of exposure analysis may be used as the basis to consider the protectiveness of a possible

environmental criterion for regulation or to judge whether an existing exposure might present risk.

A sufficient basis to support this nonlinear procedure will include data on responses that are key events[3] integral to the carcinogenic process. This means that the point of departure mostly will be from these precursor response data, e.g., hormone levels, mitogenic effects, rather than tumor incidence data.

The mode of action may have specific implications to be considered for risk potential of certain exposure scenarios. For instance, stimulus of cell growth through hormonal or other signal disruption or as a result of damage from toxicity are reversible if the exposure is for a short time since homeostasis brings a return to normal levels after cessation of exposure. Another feature of a specific exposure scenario may be the exposure of a sensitive subpopulation. If the population exposed in a particular scenario is wholly or largely composed of a subpopulation for whom evidence indicates a special sensitivity to the agent's mode of action, an adequate margin of exposure would be larger than for general population exposure.

When the mode of action information indicates that the dose-response may be adequately described by both a linear and a nonlinear approach, then the default is to present both the linear and margin of exposure analyses. An assessment may use both linear and nonlinear approaches if linearity is not plausible and nonlinearity has support, but a mode of action is not defined, or different responses are thought to result from different modes of action or a response appears to be very different at high and low doses due to influence of separate modes of action. The results may be needed for assessment of combined risk from agents with common modes of action.

A default assumption is made that cumulative dose received over a lifetime, expressed as a lifetime average daily dose, is an appropriate measure of dose. This assumes that a high dose of such an agent received over a shorter period of time is equivalent to a low dose spread over a lifetime. This is thought to be a relatively public health protective assumption and has empirical support (Monro, 1992). An example of effects of short-term, high exposure that results in subsequent cancer development is treatment of cancer patients with certain chemotherapeutic agents. An example of cancer from long-term exposure to an agent of relatively low potency is smoking. When sufficient information is available indicating that the carcinogenic mode of action supports a nonlinear dose-response approach, a different approach may be used. Such an approach includes considering the margin of exposure that exists between exposure and the point of departure from the observed data range. In these cases, short-term exposure estimates (several

[3]A "key event" is an empirically observed precursor consistent with a mode of action.

days to several months may be more appropriate than the lifetime average daily dose. In these cases both agent concentration and duration are likely to be important, because such effects are generally observed to be reversible at cessation of very short-term exposure.

1.4. CHARACTERIZATIONS

The risk characterization process first summarizes findings on hazard, dose-response, and exposure characterizations, then develops an integrative analysis of the whole risk case. It ends in a non technical Risk Characterization Summary. The Risk Characterization Summary is a presentation for risk managers who may or may not be familiar with the scientific details of cancer assessment. It also provides information for other interested readers. The initial steps in the risk characterization process are to make building blocks in the form of characterizations of the assessments of hazard, dose-response, and exposure. The individual assessments and characterizations are then integrated to arrive at risk estimates for exposure scenarios of interest. As part of the characterization process, explicit evaluations will be made of the hazard and risk potential for susceptible populations, including children (U.S EPA 1995a,b). There are two reasons for individually characterizing the hazard, dose-response, and exposure assessments. One is that they are often done by different people than those who do the integrative analyses. The second is that there is very often a lapse of time between the conduct of hazard and dose-response analyses and the conduct of exposure assessment and integrative analysis. Thus, it is necessary to capture characterizations of assessments as the assessments are done to avoid the need to go back and reconstruct them. Finally, frequently a single hazard assessment is used by several programs for several different exposure scenarios. Figure 1-2 shows the relationships of analyses. The figure does not necessarily correspond to the number of documents involved; there may be one or several. "Integrative analysis" is a generic term. At EPA, the documents of various programs that contain integrative analyses have other names such as the "Staff Paper" that discusses air quality criteria issues. In the following sections, the elements of this figure are discussed.

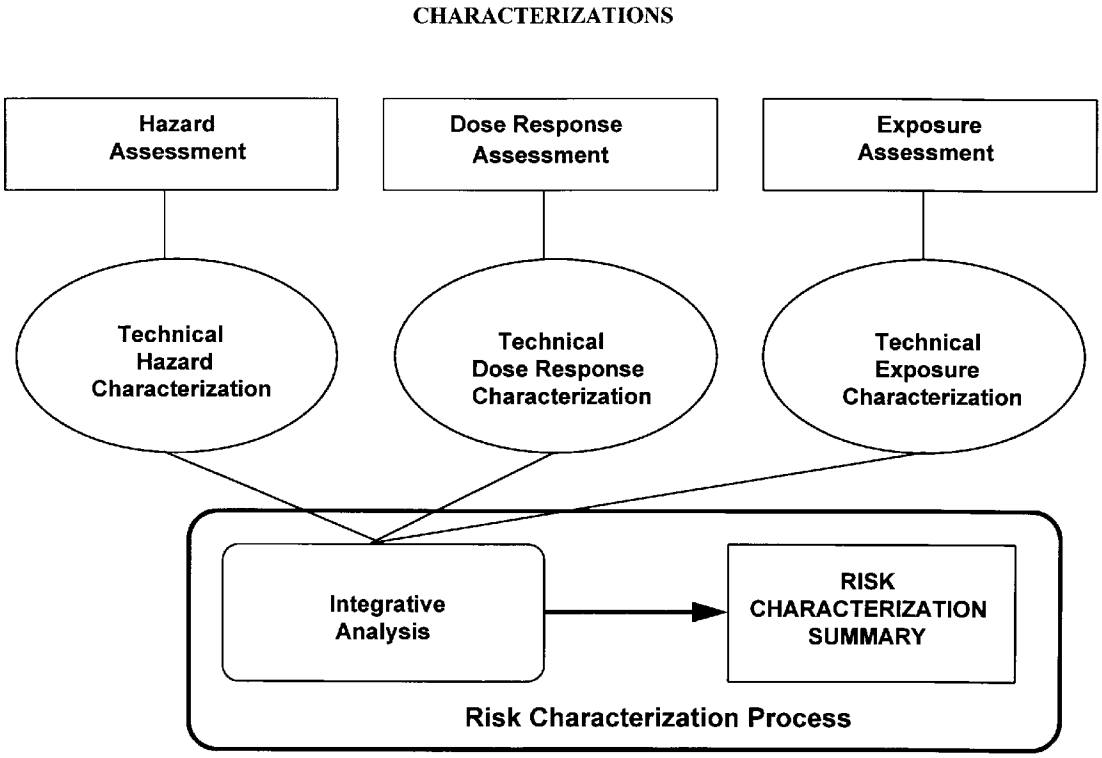

Figure 1-1. Risk Characterization

2. HAZARD ASSESSMENT

2.1. OVERVIEW OF HAZARD ASSESSMENT AND CHARACTERIZATION

2.1.1. Analyses of Data

The purpose of hazard assessment is to review and evaluate data pertinent to two questions: (1) whether an agent may pose a carcinogenic hazard to human beings and (2) under what circumstances an identified hazard may be expressed (NRC, 1994, p. 142). Hazard assessment is composed of analyses of a variety of data that may range from observations of tumor responses to analysis of structure-activity relationships. The purpose of the assessment is not simply to assemble these separate evaluations; its purpose is to construct a total case analysis examining the biological story the data reveal as a whole about carcinogenic effects, mode of action, and implications of these for human hazard and dose-response evaluation. Weight-of-evidence conclusions come from the combined strength and coherence of inferences appropriately drawn from all of the available evidence. To the extent that data permit, hazard assessment addresses the question of mode of action as both an initial step in identifying human hazard potential and as a part of considering appropriate approaches to dose-response assessment.

The topics in this chapter include analysis of tumor data, both animal and human, and analysis of other key information about properties and effects that relate to carcinogenic potential. The chapter addresses how information can be used to evaluate potential modes of action. It also provides guidance on performing a weight-of-evidence evaluation.[1]

2.1.2. Presentation of Results

Presentation of the results of hazard assessment follows Agency guidance as discussed in Section 2.7. The results are presented in a technical hazard characterization that serves as a support to later risk characterization. It includes:

- a summary of the evaluations of hazard data,
- the rationales for its conclusions, and
- an explanation of the significant strengths or limitations of the conclusions.

Another presentation feature is the use of a weight-of-evidence narrative that includes

[1]"Mode" of action is contrasted with "mechanism" of action, which implies a more detailed, molecular description of key processes and events than is meant by mode of action.

both a conclusion about the weight-of-evidence of carcinogenic potential and a summary of the data on which the conclusion rests. This narrative is a brief summary that replaces the alphanumerical classification system used in EPA's previous guidelines.

2.2. ANALYSIS OF TUMOR DATA

Evidence of carcinogenicity comes from finding tumor increases in humans or laboratory animals exposed to a given agent, or from finding tumors following exposure to structural analogues to the compound under review. The significance of observed or anticipated tumor effects is evaluated in reference to all the other key data on the agent. This section contains guidance for analyzing human and animal studies to decide whether there is an association between exposure to an agent or a structural analogue and occurrence of tumors. Note that the use of the term "tumor" here is generic, meaning malignant neoplasms or a combination of malignant and corresponding benign neoplasms.

Observation of only **benign neoplasia** may or may not have significance. Benign tumors that are not observed to progress to malignancy are assessed on a case-by-case basis. There is a range of possibilities for their overall significance. They may deserve attention because they are serious health problems even though they are not malignant; for instance, benign tumors may be a health risk because of their effect on the function of a target tissue such as the brain. They may be significant indicators of the need for further testing of an agent if they are observed in a short-term test protocol, or such an observation may add to the overall weight of evidence if the same agent causes malignancies in a long-term study. Knowledge of the mode of action associated with a benign tumor response may aid in the interpretation of other tumor responses associated with the same agent. In other cases, observation of a benign tumor response alone may have no significant health hazard implications when other sources of evidence show no suggestion of carcinogenicity.

2.2.1. Human Data

Human data may come from epidemiologic studies or case reports. Epidemiology is the study of the distributions and causes of disease within human populations. The goals of cancer epidemiology are to identify differences in cancer risk between different groups in a population or between different populations, and then to determine the extent to which these differences in risk can be attributed causally to specific exposures to exogenous or endogenous factors. Epidemiologic data are extremely useful in risk assessment because they provide direct evidence that a substance produces cancer in humans, thereby avoiding the problem of species-to-species

inference. Thus, when available human data are extensive and of good quality, they are generally preferable over animal data and should be given greater weight in hazard characterization and dose-response assessment, although both are utilized.

Null results from a single epidemiologic study cannot prove the absence of carcinogenic effects because they can arise either from being truly negative or from inadequate statistical power, inadequate design, imprecise estimates, or confounding factors. However, null results from a well-designed and well-conducted epidemiologic study that contains usable exposure data can help to define upper limits for the estimated dose of concern for human exposure if the overall weight of the evidence indicates that the agent is potentially carcinogenic in humans.

Epidemiology can also complement experimental evidence in corroborating or clarifying the carcinogenic potential of the agent in question. For example, observations from epidemiologic studies that elevated cancer incidence occurs at sites corresponding to those at which laboratory animals experience increased tumor incidence can strengthen the weight of evidence of human carcinogenicity. On the other hand, strong nonpositive epidemiologic data alone or in conjunction with compelling mechanistic information can lend support to a conclusion that animal responses may not be predictive of a human response. Furthermore, the advent of biochemical or molecular epidemiology may help improve understanding of the mechanisms of human carcinogenesis.

2.2.1.1. *Types of Studies*

The major types of cancer epidemiologic studies are analytical studies and descriptive or correlation studies. Each study type has well-known strengths and weaknesses that affect interpretation of results as summarized below (Kelsey et al., 1986; Lilienfeld and Lilienfeld, 1979; Mausner and Kramer, 1985; Rothman, 1986).

Analytical epidemiologic studies are most useful for identifying an association between human exposure and adverse health effects. Analytical study designs include case-control studies and cohort studies. In case-control studies, groups of individuals with (cases) and without (controls) a particular disease are identified and compared to determine differences in exposure. In cohort studies, a group of “exposed” and “nonexposed” individuals are identified and studied over time to determine differences in disease occurrence. Cohort studies can either be performed prospectively, or retrospectively from historical records.

Descriptive or correlation epidemiologic studies (sometimes called ecological studies) examine differences in disease rates among populations in relation to age, gender, race, and differences in temporal or environmental conditions. In general, these studies can only identify patterns or trends in disease occurrence over time or in different geographical locations, but

cannot ascertain the causal agent or degree of exposure. These studies, however, are often very useful for generating hypotheses for further research.

Biochemical or molecular epidemiologic studies are studies in which laboratory methods are incorporated in analytical investigations. The application of techniques for measuring cellular and molecular alterations due to exposure to specific environmental agents may allow conclusions to be drawn about the mechanisms of carcinogenesis. The use of biological biomarkers in epidemiology may improve assessment of exposure and internal dose.

Case reports describe a particular effect in an individual or group of individuals who were exposed to a substance. These reports are often anecdotal or highly selected in nature and are of limited use for hazard assessment. However, reports of cancer cases can identify associations, particularly when there are unique features such as an association with an uncommon tumor (e.g., vinyl chloride and angiosarcoma or diethylstilbestrol and clear-cell carcinoma of the vagina).

2.2.1.2. *Criteria for Assessing Adequacy of Epidemiologic Studies*

Criteria for assessing the adequacy of epidemiologic studies are well recognized. Characteristics that are desirable in these studies include (1) clear articulation of study objectives or hypothesis; (2) proper selection and characterization of the exposed and control groups; (3) adequate characterization of exposure; (4) sufficient length of follow-up for disease occurrence; (5) valid ascertainment of the causes of cancer morbidity and mortality; (6) proper consideration of bias and confounding factors; (7) adequate sample size to detect an effect; (8) clear, well-documented, and appropriate methodology for data collection and analysis; (9) adequate response rate and methodology for handling missing data; and (10) complete and clear documentation of results. Ideally, these conditions should be satisfied, where appropriate, but rarely can a study meet all of them. No single criterion determines the overall adequacy of a study. The following discussions highlight the major factors included in an analysis of epidemiologic studies.

Population Issues

The ideal comparison would be between two populations that differ only in exposure to the agent in question. Because this is seldom the case, it is important to identify sources of bias inherent in a study's design or data collection methods. Bias can arise from several sources, including noncomparability between populations of factors such as general health (McMichael, 1976), diet, lifestyle, or geographic location; differences in the way case and control individuals recall past events; differences in data collection that result in unequal ascertainment of health effects in the populations; and unequal follow-up of individuals. Both acceptance of studies for

assessment and judgment of their strengths or weaknesses depend on identifying their sources of bias and the effects on study results.

Exposure Issues

For epidemiologic data to be useful in determining whether there is an association between health effects and exposure to an agent, there must be adequate characterization of exposure information. In general, greater weight should be given to studies with more precise and specific exposure estimates.

Questions to address about exposure are: What can one reliably conclude about the level, duration, route, and frequency of exposure of individuals in one population as compared with another? How sensitive are study results to uncertainties in these parameters?

Actual exposure measurements are not available for many retrospective studies. Therefore, surrogates are often used to reconstruct exposure parameters. These may involve attributing exposures to job classifications in a workplace or to broader occupational or geographic groupings. Use of surrogates carries a potential for misclassification in that individuals may be placed in an incorrect exposure group. Misclassification generally leads to reduced ability of a study to detect differences between study and referent populations.

When either current or historical monitoring data are available, the exposure evaluation includes consideration of the error bounds of the monitoring and analytic methods and whether the data are from routine or accidental exposures. The potentials for misclassification and measurement errors are amenable to both qualitative and quantitative analysis. These are essential analyses for judging a study's results because exposure estimation is the most critical part of a retrospective study.

Biological markers potentially offer excellent measures of exposure (Hulka and Margolin, 1992; Peto and Darby, 1994). Validated markers of exposure such as alkylated hemoglobin from exposure to ethylene oxide (van Sittert et al., 1985) or urinary arsenic (Enterline et al., 1987) can greatly improve estimates of dose. Markers closely identified with effects promise to greatly increase the ability of studies to distinguish real effects from bias at low levels of relative risk between populations (Taylor et al., 1994; Biggs et al., 1993) and to resolve problems of confounding risk factors.

Confounding Factors

Because epidemiologic studies are mostly observational, it is not possible to guarantee the

control of confounding variables, which may affect the study outcome. A confounding variable is a risk factor, independent of the putative agent, that is distributed unequally among the exposed and unexposed populations (e.g., smoking habits, lifestyle). Adjustment for possible confounding factors can occur either in the design of the study (e.g., matching on critical factors) or in the statistical analysis of the results. The influence of a potential confounding factor is limited by the effect of the exposure of interest. For example, a twofold effect of an exposure requires that the confounder effect be at least as big. The latter may not be possible owing to the presentation of the data or because needed information was not collected during the study. In this case, indirect comparisons may be possible. For example, in the absence of data on smoking status among individuals in the study population, an examination of the possible contribution of cigarette smoking to increased lung cancer risk may be based on information from other sources such as the American Cancer Society's longitudinal studies (Hammand, 1966; Garfinkel and Silverberg, 1991). The effectiveness of adjustments contributes to the ability to draw inferences from a study.

Different studies involving exposure to an agent may have different confounding factors. If consistent increases in cancer risk are observed across a collection of studies with different confounding factors, the inference that the agent under investigation was the etiologic factor is strengthened, even though complete adjustment for confounding factors cannot be made and no single study supports a strong inference.

It also may be the case that the agent of interest is a risk factor in conjunction with another agent. This relationship may be revealed in a collection of studies such as in the case of asbestos exposure and smoking.

Sensitivity

Sensitivity, or the ability of a study to detect real effects, is a function of several factors. Greater size of the study population(s) (sample size) increases sensitivity, as does greater exposure (levels and duration) of the population members. Because of the often long latency period in cancer development, sensitivity also depends on whether adequate time has elapsed since exposure began for effects to occur. A unique feature that can be ascribed to the effects of a particular agent (such as a tumor type that is seen only rarely in the absence of the agent) can increase sensitivity by permitting separation of bias and confounding factors from real effects. Similarly, a biomarker particular to the agent can permit these distinctions. Statistical re-analyses of data, particularly an examination of different exposure indices, can give insight on potential exposure-response relationships. These are all factors to explore in statistical analysis of the data.

Statistical Considerations

The analysis applies appropriate statistical methods to ascertain whether or not there is any significant association between exposure and effects. A description of the method or methods should include the reasons for their selection. Statistical analyses of the potential effects of bias or confounding factors are part of addressing the significance of an association, or lack of one, and whether a study is able to detect any effect.

The analysis augments examination of the results for the whole population with exploration of the results for groups with comparatively greater exposure or time since first exposure. This may support identifying an association or establishing a dose-response trend. When studies show no association, such exploration may apply to determining an upper limit on potential human risk for consideration alongside results of animal tumor effects studies.

Combining Statistical Evidence Across Studies

Meta-analysis is a means of comparing and synthesizing studies dealing with similar health effects and risk factors. It is intended to introduce consistency and comprehensiveness into what otherwise might be a more subjective review of the literature. When utilized appropriately, meta-analysis can enhance understanding of associations between sources and their effects that may not be apparent from examination of epidemiologic studies individually. Whether to conduct a meta-analysis depends on several issues. These include the importance of formally examining sources of heterogeneity, the refinement of the estimate of the magnitude of an effect, and the need for information beyond that provided by individual studies or a narrative review. Meta-analysis may not be useful in some circumstances. These include when the relationship between exposure and disease is obvious without a more formal analysis; when there are only a few studies of the key health outcomes; when there is insufficient information from available studies related to disease, risk estimate, or exposure classification; or when there are substantial confounding or other biases that cannot be adjusted for in the analysis (Blair et al., 1995; Greenland, 1987; Peto, 1992).

2.2.1.3. *Criteria for Causality*

A causal interpretation is enhanced for studies to the extent that they meet the criteria described below. None of the criteria is conclusive by itself, and the only criterion that is essential is the temporal relationship. These criteria are modeled after those developed by Bradford Hill in the examination of cigarette smoking and lung cancer (Rothman, 1986), and they need to be interpreted in the light of all other information on the agent being assessed.

- Temporal relationship: The development of cancers requires certain latency periods, and while latency periods vary, existence of such periods is generally acknowledged. Thus, the disease has to occur within a biologically reasonable time after initial exposure. This feature must be present if causality is to be considered.
- Consistency: Associations occur in several independent studies of a similar exposure in different populations, or associations occur consistently for different subgroups in the same study. This feature usually constitutes strong evidence for a causal interpretation when the same bias or confounding is not also duplicated across studies.
- Magnitude of the association: A causal relationship is more credible when the risk estimate is large and precise (narrow confidence intervals).
- Biological gradient: The risk ratio (i.e., the ratio of the risk of disease or death among the exposed to the risk of the unexposed) increases with increasing exposure or dose. Statistical significance is important, and a strong dose-response relationship across several categories of exposure, latency, and duration is supportive for causality, given that confounding is unlikely to be correlated with exposure. The absence of a dose-response relationship, however, is not by itself evidence against a causal relationship.
- Specificity of the association: The likelihood of a causal interpretation is increased if an exposure produces a specific effect (one or more tumor types also found in other studies) or if a given effect has a unique exposure.
- Biological plausibility: The association makes sense in terms of biological knowledge. Information is considered from animal toxicology, toxicokinetics, structure-activity relationship analysis, and short-term studies of the agent's influence on events in the carcinogenic process considered.
- Coherence: The cause-and-effect interpretation is in logical agreement with what is known about the natural history and biology of the disease, i.e., the entire body of knowledge about the agent.

2.2.1.4. *Assessment of Evidence of Carcinogenicity from Human Data*

In the evaluation of carcinogenicity based on epidemiologic studies, it is necessary to critically evaluate each study for confidence in findings and conclusions as discussed under Section 2.2.1.2. All studies that are properly conducted, whether yielding positive or null results,

or even suggesting protective carcinogenic effects, should be considered in assessing the totality of the human evidence. Although a single study may be indicative of a cause-effect relationship, confidence in inferring a causal relationship is increased when several independent studies are concordant in showing the association, when the association is strong, and when other criteria for causality are also met. Conclusions about the overall evidence for carcinogenicity from available studies in humans should be summarized along with a discussion of strengths or limitations of the conclusions.

2.2.2. Animal Data

Various whole-animal test systems are currently used or are under development for evaluating potential carcinogenicity. Cancer studies involving chronic exposure for most of the lifespan of an animal are generally accepted for evaluation of tumor effects (Tomatis et al., 1989; Rall, 1991; Allen et al., 1988; but see Ames and Gold, 1990). Other studies of special design are useful for observing formation of preneoplastic lesions or tumors or investigating specific modes of action. Their applicability is made on a case-by-case basis.

2.2.2.1. *Long-Term Carcinogenicity Studies*

The objective of long-term carcinogenesis bioassays is to determine the potential carcinogenic hazard and dose-response relationships of the test agent. Carcinogenicity rodent studies are designed to examine the production of tumors as well as preneoplastic lesions and other indications of chronic toxicity that may provide evidence of treatment-related effects and insights into the way the test agent produces tumors. Current standardized carcinogenicity studies in rodents test at least 50 animals per sex per dose group in each of three treatment groups and in a concurrent control group, usually for 18 to 24 months, depending on the rodent species tested (OECD, 1981; U.S. EPA, 1983a-c). The high dose in long-term studies is generally selected to provide the maximum ability to detect treatment-related carcinogenic effects while not compromising the outcome of the study through excessive toxicity or inducing inappropriate toxicokinetics (e.g., overwhelming absorption or detoxification mechanisms). The purpose of two or more lower doses is to provide some information on the shape of the dose-response curve. Similar protocols have been and continue to be used by many laboratories worldwide.

All available studies of tumor effects in whole animals are considered, at least preliminarily. The analysis discards studies judged to be wholly inadequate in protocol, conduct, or results. Criteria for the technical adequacy of animal carcinogenicity studies have been published and should be used as guidance to judge the acceptability of individual studies (NTP,

1984; OSTP, 1985). Care is taken to include studies that provide some evidence bearing on carcinogenicity or that help interpret effects noted in other studies, even if they have some limitations of protocol or conduct. Such limited, but not wholly inadequate, studies can contribute as their deficiencies permit. The findings of long-term rodent bioassays are always interpreted in conjunction with results of prechronic studies along with metabolism toxicokinetic metabolism studies and other pertinent information, if available. Evaluation of tumor effects requires consideration of both biological and statistical significance of the findings (Haseman, 1984, 1985, 1990, 1995). The following sections highlight the major issues in the evaluation of long-term carcinogenicity studies.

Dosing Issues

Among the many criteria for technical adequacy of animal carcinogenicity studies is the appropriateness of dose selection. The selection of doses for chronic bioassays requires scientific judgments and must be based on sound toxicologic principles. Dose selection should be made on the basis of relevant toxicologic information from prechronic, mechanistic, and toxicokinetic and mechanistic studies. How well the dose selection is made can be evaluated only after the completion of the bioassay. A scientific rationale for dose selection should be clearly articulated (ILSI, 1997).

In order to obtain the most relevant information from a long-term carcinogenicity study, it is important to maximize exposure conditions to the test material. At the same time, there is a need for caution in using excessive high-dose levels that would confound the interpretation of study results to humans. The middle and lowest doses should be selected to characterize the shape of the dose-response curve as much as possible. It is important that the doses are adequately spaced so that the study would provide relevant dose-response data for assessing human hazard and risk. If the testing of potential carcinogenicity is being combined with an evaluation of noncancer chronic toxicity, the study should be designed to include one dose that does not elicit adverse effects.

With regard to the appropriateness of the high dose, an adequate high dose would be one that produces some toxic effects without either unduly affecting mortality from effects other than cancer or producing significant adverse effects on the nutrition and health of the test animals (OECD, 1981; NRC, 1993b). If the test agent does not appear to cause any specific target organ toxicity or perturbation of physiological function, an adequate high dose would be one that causes no more than 5%-10% reduction of body weight gain over the lifespan of the animals. The high dose would be considered inadequate if no toxicity is observed. On the other hand, significant

increases in mortality from effects other than cancer generally indicate that an adequate high dose has been exceeded. Other signs of treatment-related toxicity associated with an excessive high dose may include the following: (a) reduction of body weight gain greater than 10%, (b) significant increases in abnormal behavioral and clinical signs, (c) significant changes in hematology or clinical chemistry, (d) saturation of absorption and detoxification mechanisms, or (e) marked changes in organ weight, morphology, and histopathology. It should be noted that practical upper limits have been established to avoid the use of excessively high doses in long-term carcinogenicity studies of environmental chemicals (e.g., 5% of the test substance in the feed for dietary studies or 1 g/kg of body weight for oral gavage studies [OECD, 1981]).

For dietary studies, weight gain reductions should be evaluated as to whether there is a palatability problem or an issue with food efficiency; certainly, the latter is a toxic manifestation. In the case of inhalation studies with respirable particles, evidence of impairment of normal clearance of particles from the lung should be considered along with other signs of toxicity to the respiratory airways to determine whether the high exposure concentration has been appropriately selected. For dermal studies, evidence of skin irritation may indicate that an adequate high dose has been reached (U.S. EPA, 1989d).

Interpretation of carcinogenicity study results is profoundly affected by study exposure conditions, especially by inappropriate dose selection. This is particularly important in studies that are nonpositive for carcinogenicity, since failure to reach a sufficient dose reduces the sensitivity of the studies. A lack of tumorigenic responses at exposure levels that cause significant impairment of animal survival may also not be acceptable. In addition, overt toxicity or inappropriate toxicokinetics due to excessively high doses may result in tumor effects that are secondary to the toxicity rather than directly attributable to the agent.

There are several possible outcomes regarding the study interpretation of the significance and relevance of tumorigenic effects associated with exposure or dose levels below, at, or above an adequate high dose. General guidance is given here that should not be taken as prescriptive; for each case, the information at hand is evaluated and a rationale should be given for the position taken.

- Adequate high dose: If an adequate high dose has been utilized, tumor effects are judged positive or negative depending on the presence or absence of significant tumor incidence increases, respectively.
- Excessive high dose: If toxicity or mortality is excessive at the high dose, interpretation depends on the finding of tumors or not.

(a) Studies that show tumor effects only at excessive doses may be compromised and may or may not carry weight, depending on the interpretation in the context of other study results and other lines of evidence. Results of such studies, however, are generally not considered suitable for dose-response extrapolation if it is determined that the mode(s) of action underlying the tumorigenic responses at high doses are not operative at lower doses.

(b) Studies that show tumors at lower doses, even though the high dose is excessive and may be discounted, should be evaluated on their own merits.

(c) If a study does not show an increase in tumor incidence at a toxic high dose and appropriately spaced lower doses are used without such toxicity or tumors, the study is generally judged as negative for carcinogenicity.

- Inadequate high dose: Studies of inadequate sensitivity where an adequate high dose has not been reached may be used to bound the dose range where carcinogenic effects might be expected.

Statistical Considerations

The main aim of statistical evaluation is to determine whether exposure to the test agent is associated with an increase of tumor development. Statistical analysis of a long-term study should be performed for each tumor type separately. The incidence of benign and malignant lesions of the same cell type, usually within a single tissue or organ, are considered separately and are combined when scientifically defensible (McConnell et al., 1986).

Trend tests and pairwise comparison tests are the recommended tests for determining whether chance, rather than a treatment-related effect, is a plausible explanation for an apparent increase in tumor incidence. A trend test such as the Cochran-Armitage test (Snedecor and Cochran, 1967) asks whether the results in all dose groups together increase as dose increases. A pairwise comparison test such as the Fisher exact test (Fisher, 1950) asks whether an incidence in one dose group is increased over the control group. By convention, for both tests a statistically significant comparison is one for which $p < 0.05$ that the increased incidence is due to chance. Significance in either kind of test is sufficient to reject the hypothesis that chance accounts for the result. A statistically significant response may or may not be biologically significant and vice versa. The selection of a significance level is a policy choice based on a trade-off between the risks of false positives and false negatives. A significance level of greater or less than 5% is examined to see if it confirms other scientific information. When the assessment departs from a

simple 5% level, this should be highlighted in the risk characterization. A two-tailed test or a one-tailed test can be used. In either case a rationale is provided.

Considerations of multiple comparisons should also be taken into account. Haseman (1983) analyzes typical animal bioassays testing both sexes of two species and concludes that, because of multiple comparisons, a single tumor increase for a species-sex-site combination that is statistically significant at the 1% level for common tumors or 5% for rare tumors corresponds to a 7%-8% significance level for the study as a whole. Therefore, animal bioassays presenting only one significant result that falls short of the 1% level for a common tumor must be treated with caution.

Concurrent and Historical Controls

The standard for determining statistical significance of tumor incidence comes from a comparison of tumors in dosed animals as compared with concurrent control animals. Additional insights about both statistical and biological significance can come from an examination of historical control data (Tarone, 1982; Haseman, 1995). Historical control data can add to the analysis, particularly by enabling identification of uncommon tumor types or high spontaneous incidence of a tumor in a given animal strain. Identification of common or uncommon situations prompts further thought about the meaning of the response in the current study in context with other observations in animal studies and with other evidence about the carcinogenic potential of the agent. These other sources of information may reinforce or weaken the significance given to the response in the hazard assessment. Caution should be exercised in simply looking at the ranges of historical responses because the range ignores differences in survival of animals among studies and is related to the number of studies in the database.

In analyzing results for uncommon tumors in a treated group that are not statistically significant in comparison to concurrent controls, the analyst can use the experience of historical controls to conclude that the result is in fact unlikely to be due to chance. In analyzing results for common tumors, a different set of considerations comes into play. Generally speaking, statistically significant increases in tumors should not be discounted simply because incidence rates in the treated groups are within the range of historical controls or because incidence rates in the concurrent controls are somewhat lower than average. Random assignment of animals to groups and proper statistical procedures provide assurance that statistically significant results are unlikely to be due to chance alone. However, caution should be used in interpreting results that are barely statistically significant or in which incidence rates in concurrent controls are unusually low in comparison with historical controls.

In cases where there may be reason to discount the biological relevance to humans of increases in common animal tumors, such considerations should be weighed on their own merits and clearly distinguished from statistical concerns.

When historical control data are used, the discussion needs to address several issues that affect comparability of historical and concurrent control data. Among these issues are the following: genetic drift in the laboratory strains, differences in pathology examination at different times and in different laboratories (e.g., in criteria for evaluating lesions; variations in the techniques for preparation or reading of tissue samples among laboratories), and comparability of animals from different suppliers. The most relevant historical data come from the same laboratory and same supplier, gathered within 2 or 3 years one way or the other of the study under review; other data should be used only with extreme caution.

Assessment of Evidence of Carcinogenicity from Long-Term Animal Studies

In general, observation of tumor effects under different circumstances lends support to the significance of the findings for animal carcinogenicity. Significance is a function of the number of factors present and, for a factor such as malignancy, the severity of the observed pathology. The following observations add significance to the tumor findings:

- uncommon tumor types;
- tumors at multiple sites;
- tumors by more than one route of administration;
- tumors in multiple species, strains, or both sexes;
- progression of lesions from preneoplastic to benign to malignant;
- reduced latency of neoplastic lesions;
- metastases;
- unusual magnitude of tumor response;
- proportion of malignant tumors; and
- dose-related increases.

These guidelines adopt the science policy position that tumor findings in animals indicate that an agent may produce such effects in humans. Moreover, the absence of tumor findings in well-conducted, long-term animal studies in at least two species provides reasonable assurance that an agent may not be a carcinogenic concern for humans. Each of these is a default assumption that may be adopted, when appropriate, after evaluation of tumor data and other key

evidence.

Site Concordance

Site concordance of tumor effects between animals and humans is an issue to be considered in each case. Thus far, there is evidence that growth control mechanisms at the level of the cell are homologous among mammals, but there is no evidence that these mechanisms are site concordant. Moreover, agents observed to produce tumors in both humans and animals have produced tumors either at the same (e.g., vinyl chloride) or different sites (e.g., benzene) (NRC, 1994). Hence, site concordance is not assumed a priori. On the other hand, certain processes with consequences for particular tissue sites (e.g., disruption of thyroid function) may lead to an anticipation of site concordance.

2.2.2.2. *Perinatal Carcinogenicity Studies*

The objective of perinatal carcinogenesis studies is to determine the carcinogenic potential and dose-response relationships of the test agent in the developing organism. Some investigators have postulated that the age of initial exposure to a chemical carcinogen may influence the carcinogenic response (Vesselinovitch et al., 1979; Rice, 1979; McConnell, 1992). Current standardized long-term carcinogenesis bioassays generally begin dosing animals at 6-8 weeks of age and continue dosing for the lifespan of the animal (18-24 months). This protocol has been modified in some cases to investigate the potential of the test agent to induce transplacental carcinogenesis or to investigate the potential differences following perinatal and adult exposures; but currently there is not a standardized protocol for testing agents for carcinogenic effects following prenatal or early postnatal exposure.

Several cancer bioassay studies have compared adult and perinatal exposures (see McConnell, 1992; U.S. EPA, 1996a). A review of these reveals that perinatal exposure rarely identifies carcinogens that are not found in standard animal bioassays. Exposure that is perinatal sometimes slightly increases the incidence of a given type of tumor. The increase may reflect an increased length of exposure and a higher dose for the developing organism relative to the adult, or an increase in sensitivity in some cases. Additionally, exposure that is perinatal through adulthood sometimes reduces the latency period for tumors to develop in the growing organism (U.S. EPA, 1996a).

Because the perinatal exposure studies done to date provide only marginal additions to knowledge as compared with standard bioassay protocols, EPA evaluates the need for such a study agent-by-agent (U.S. EPA, 1997a,b). Perinatal study data analysis follows the principles discussed above for evaluating other long-term carcinogenicity studies. When differences in

responses in perinatal animals compared to adult animals suggest an increased susceptibility or sensitivity of perinatal or postnatal animals, such as the ones below, a separate evaluation of the response is prepared:

- a difference in dose-response relationship
- presence of different tumor types
- an earlier onset of tumors
- an increase in the incidence of tumors

An illustrative case study appears in Appendix E.

2.2.2.3. *Other Studies*

Various intermediate-term studies often use protocols that screen for carcinogenic or preneoplastic effects, sometimes in a single tissue. Some involve the development of various proliferative lesions, like foci of alteration in the liver (Goldsworthy et al., 1986). Others use tumor endpoints, like the induction of lung adenomas in the sensitive strain A mouse (Maronpot et al., 1986) or tumor induction in initiation-promotion studies using various organs such as the bladder, intestine, liver, lung, mammary gland, and thyroid (Ito et al., 1992). In these tests, the selected tissue is, in a sense, the test system rather than the whole animal. Important information concerning the steps in the carcinogenic process and mode of action can be obtained from "start/stop" experiments. In these protocols, an agent is given for a period of time to induce particular lesions or effects, then stopped to evaluate the progression or reversibility of processes (Todd, 1986; Marsman and Popp, 1994).

Assays in genetically engineered rodents may provide insight into the chemical and gene interactions involved in carcinogenesis (Tennant et al., 1995). These mechanistically based approaches involve activated oncogenes that are introduced (transgenic) or tumor suppressor genes that are deleted (knocked out). If appropriate genes are selected, not only may these systems provide information on mechanisms, but the rodents typically show tumor development earlier than the standard bioassay. Transgenic mutagenesis assays also represent a mechanistic approach for assessing the mutagenic properties of agents as well as developing quantitative linkages between exposure, internal dose, and mutation related to tumor induction (Morrison and Ashby, 1994; Sisk et al., 1994; Hayward et al., 1995). These systems use a stable genomic integration of a lambda shuttle vector that carries a *lacI* target gene and a *lacZ* reporter gene.

The support that these studies give to a determination of carcinogenicity rests on their contribution to the consistency of other evidence about an agent. For instance, benzoyl peroxide

has promoter activity on the skin, but the overall evidence may be less supportive (Kraus et al., 1995). These studies also may contribute information about mode of action. One needs to recognize the limitations of these experimental protocols such as short duration, limited histology, lack of complete development of tumors, or experimental manipulation of the carcinogenic process that may limit their contribution to the overall assessment. Generally, their results are appropriate as aids in the assessment for interpreting other toxicological evidence (e.g., rodent chronic bioassays), especially regarding potential modes of action. With sufficient validation, these studies may partially or wholly replace chronic bioassays in the future (Tennant et al., 1995).

2.2.3. Structural Analogue Data

For some chemical classes, there is significant information available on the carcinogenicity of analogues, largely in rodent bioassays. Analogue effects are instructive in investigating carcinogenic potential of an agent as well as identifying potential target organs, exposures associated with effects, and potential functional class effects or modes of action. All appropriate studies are included and analyzed, whether indicative of a positive effect or not. Evaluation includes tests in various animal species, strains, and sexes; with different routes of administration; and at various doses, as data are available. Confidence in conclusions is a function of how similar the analogues are to the agent under review in structure, metabolism, and biological activity. This confidence needs to be considered to ensure a balanced position.

2.3. ANALYSIS OF OTHER KEY DATA

The physical, chemical, and structural properties of an agent, as well as data on endpoints that are thought to be critical elements of the carcinogenic process, provide valuable insights into the likelihood of human cancer risk. The following sections provide guidance for analyses of these data.

2.3.1. Physicochemical Properties

Physicochemical properties affect an agent's absorption, tissue distribution (bioavailability), biotransformation, and degradation in the body and are important determinants of hazard potential (and dose-response analysis). Properties to analyze include, but are not limited to, the following: molecular weight, size, and shape; valence state; physical state (gas, liquid, solid); water or lipid solubility, which can influence retention and tissue distribution; and potential for chemical degradation or stabilization in the body.

An agent's potential for chemical reaction with cellular components, particularly with

DNA and proteins, is also important. The agent's molecular size and shape, electrophilicity, and charge distribution are considered in order to decide whether they would facilitate such reactions.

2.3.2. Structure-Activity Relationships

Structure-activity relationship (SAR) analyses and models can be used to predict molecular properties, surrogate biological endpoints, and carcinogenicity. Overall, these analyses provide valuable initial information on agents, may strengthen or weaken concern, and are part of the weight of evidence.

Currently, SAR analysis is most useful for chemicals and metabolites that are believed to initiate carcinogenesis through covalent interaction with DNA (i.e., DNA-reactive, mutagenic, electrophilic, or proelectrophilic chemicals) (Ashby and Tennant, 1991). For organic chemicals, the predictive capability of SAR analysis combined with other toxicity information has been demonstrated (Ashby and Tennant, 1994). The following parameters are useful in comparing an agent to its structural analogues and congeners that produce tumors and affect related biological processes such as receptor binding and activation, mutagenicity, and general toxicity (Woo and Arcos, 1989):

- nature and reactivity of the electrophilic moiety or moieties present;
- potential to form electrophilic reactive intermediate(s) through chemical, photochemical, or metabolic activation;
- contribution of the carrier molecule to which the electrophilic moiety(ies) is attached;
- physicochemical properties (e.g., physical state, solubility, octanol-water partition coefficient, half-life in aqueous solution);
- structural and substructural features (e.g., electronic, stearic, molecular geometric);
- metabolic pattern (e.g., metabolic pathways and activation and detoxification ratio); and
- possible exposure route(s) of the agent.

Suitable SAR analysis of non-DNA-reactive chemicals and of DNA-reactive chemicals that do not appear to bind covalently to DNA requires knowledge or postulation of the probable mode(s) of action of closely related carcinogenic structural analogues (e.g., receptor-mediated, cytotoxicity-related). Examination of the physicochemical and biochemical properties of the agent may then provide the rest of the information needed in order to make an assessment of the likelihood of the agent's activity by that mode of action.

2.3.3. Comparative Metabolism and Toxicokinetics

Studies of the absorption, distribution, biotransformation, and excretion of agents permit comparisons among species to assist in determining the implications of animal responses for human hazard assessment, supporting identification of active metabolites, identifying changes in distribution and metabolic pathway or pathways over a dose range, and making comparisons among different routes of exposure.

If extensive data are available (e.g., blood/tissue partition coefficients and pertinent physiological parameters of the species of interest), physiologically based pharmacokinetic models can be constructed to assist in a determination of tissue dosimetry, species-to-species extrapolation of dose, and route-to-route extrapolation (Connolly and Andersen, 1991; see Section 3.2.2). If it is not contrary to available data, it is assumed as a default that toxicokinetic and metabolic processes are qualitatively comparable between species. Discussion of the defaults regarding quantitative comparison and their modifications appears in Chapter 3.

The *qualitative* question of whether an agent is absorbed by a particular route of exposure is important for weight-of-evidence classification, discussed in Section 2.7.1. Decisions whether route of exposure is a limiting factor on expression of any hazard, in that absorption does not occur by a route, are based on studies in which effects of the agent, or its structural analogues, have been observed by different routes, on physical-chemical properties, or on toxicokinetics studies.

Adequate metabolism and pharmacokinetic data can be applied toward the following as data permit. Confidence in conclusions is enhanced when in vivo data are available.

- Identifying metabolites and reactive intermediates of metabolism and determining whether one or more of these intermediates are likely to be responsible for the observed effects. This information on the reactive intermediates will appropriately focus SAR analysis, analysis of potential modes of action, and estimation of internal dose in dose-response assessment (D'Souza et al., 1987; Krewski et al., 1987).
- Identifying and comparing the relative activities of metabolic pathways in animals with those in humans as well as different ages. This analysis can provide insights for extrapolating results of animal studies to humans.
- Describing anticipated distribution within the body and possibly identifying target organs. Use of water solubility, molecular weight, and structure analysis can support qualitative inferences about anticipated distribution and excretion. In addition, describing whether the agent or metabolite of concern will be excreted rapidly or

slowly or will be stored in a particular tissue or tissues to be mobilized later can identify issues in comparing species and formulating dose-response assessment approaches.

- Identifying changes in toxicokinetics and metabolic pathways with increases in dose. These changes may result in important differences in disposition of the agent or its generation of active forms of the agent between high and low dose levels. These studies play an important role in providing a rationale for dose selection in carcinogenicity studies.
- Identifying and comparing metabolic process differences by age, sex, or other characteristic so that sensitive subpopulations can be recognized. For example, metabolic capacity with respect to P450 enzymes in newborn children is extremely limited compared to adults, so that a requirement for metabolic activation of a carcinogen will limit its effect in young, whereas a requirement for metabolic deactivation will result in increased sensitivity of this subpopulation (Cresteil, 1998). A variety of changes in toxicokinetics and physiology occur from fetal to post-weaning, to young child. Any of these may make a difference to risk (Renwick, 1998)
- Determining bioavailability via different routes of exposure by analyzing uptake processes under various exposure conditions. This analysis supports identification of hazards for untested routes. In addition, use of physicochemical data (e.g., octanol-water partition coefficient information) can support an inference about the likelihood of dermal absorption (Flynn, 1990).

In all of these areas, attempts are made to clarify and describe as much as possible the variability to be expected because of differences in species, sex, age, and route of exposure. The analysis takes into account the presence of subpopulations of individuals who are particularly vulnerable to the effects of an agent because of toxicokinetic or metabolic differences (genetically or environmentally determined) (Bois et al., 1995), and is a special emphasis for assessment of risks to children.

2.3.4. Toxicological and Clinical Findings

Toxicological findings in experimental animals and clinical observations in humans are an important resource to the cancer hazard assessment. Such findings provide information on physiological effects and effects on enzymes, hormones, and other important macromolecules, as

well as on target organs for toxicity. Given that the cancer process represents defects in terminal differentiation, growth control, and cell death, developmental studies of agents may provide an understanding of the activity of an agent that carries over to cancer assessment. Toxicity studies in animals by different routes of administration support comparison of absorption and metabolism by those routes. Data on human variability in standard clinical tests may provide insight into the range of human sensitivity and common mechanisms to agents that affect the tested parameters.

2.3.5. Events Relevant to Mode of Carcinogenic Action

Information on the biochemical and biological changes that precede tumor development (which includes but is not limited to mutagenesis, increased cell proliferation, inhibition of programmed cell death, and receptor activation) may provide important information in determining whether a cancer hazard exists and may help inform the dose-response relationship below the range of observable tumor response. Because cancer is the result of a series of genetic defects in genes controlling cell growth, division, and differentiation (Vogelstein et al., 1988), the ability of an agent to affect genes or gene expression is of obvious importance in evaluating its influence on the carcinogenic process. Initial and key questions to examine are: Does the agent (or its metabolite) interact directly with and mutate DNA to bring about changes in gene expression? Does the agent bring about effects on gene expression via other processes? Furthermore, carcinogenesis involves a complex series and interplay of events that alter the signals a cell receives from its extracellular environment to promote growth. Many, but not all, mutagens are carcinogens, and some, but not all, agents that induce cell proliferation lead to tumor development. Thus, understanding the range of key influences that the chemical may have on the carcinogenic process is essential for evaluating mode of action. Endpoints that provide insight into an agent's ability to alter genes and gene expression and other features of an agent's potential mode of carcinogenic action are discussed below.

2.3.5.1. *Direct DNA Reactive Effects*

It is well known that many carcinogens are electrophiles that interact with DNA, resulting in DNA adducts and breakage (referred to in these guidelines as direct DNA effects). Following DNA replication, these DNA lesions can be converted into mutations and stable cytogenetic alterations, which then may initiate and contribute to the carcinogenic process (Shelby and Zeiger, 1990; Tinwell and Ashby, 1991). Thus, studies of mutations and other genetic lesions continue to be important in the assessment of potential human cancer hazard and in the understanding of an agent's mode of carcinogenic action. EPA has published testing guidelines for detecting the

ability of an agent to damage DNA and produce mutations and chromosomal aberrations. Briefly, standard tests for gene mutations in bacteria and mammalian cells in vitro and in vivo, and for structural chromosomal aberrations in vitro and in vivo are important examples of relevant methods. New molecular approaches such as mouse mutations and cancer transgenic models are providing a means to examine mutation at tissue sites where the tumor response is observed (Heddle and Swiger, 1996). Additionally, continued improvements in fluorescent-based chromosome staining methods (FISH, fluorescent in situ hybridization) will allow the detection of specific chromosomal abnormalities in relevant target tissues (Tucker and Preston, 1998).

Endpoints indicative of DNA damage but not measures of mutation per se, such as DNA adducts or strand breakage, can be detected in relevant target tissues and thus contribute to evaluating an agent's mutagenic potential. Evidence of chemical-specific DNA adducts (e.g., reactions at oxygen sites in DNA bases or with ring nitrogens of guanine and adenine) provides information on a mutagen's ability to directly interact with DNA (La and Swenberg, 1996). It should be noted that an increase in DNA binding shown with a radioactive label incorporated in the chemical (e.g., C^{14}) may reflect a direct DNA reactive mechanism, but needs to be examined because the label may reflect reuse of C^{14} in the synthesis of DNA rather than binding. Some planar molecules (e.g., 9-aminoacridine) intercalate between base pairs of DNA, which results in a physical distortion in DNA that may lead to mutations when DNA replicates. As discussed below, some carcinogens do not interact directly with DNA, but can produce increases in endogenous levels of DNA adducts (e.g., 8-hydroxyguanine) by indirect mechanisms.

2.3.5.2. *Indirect DNA Effects or Other Effects on Genes/Gene Expression*

Although some carcinogens may result in an elevation of mutations or cytogenetic anomalies as detected in standard assays, they may do so by indirect mechanisms. These effects may be brought about by chemical-cell interactions rather than the chemical (or its metabolite) directly interacting with DNA. An increase in mutations might be due to cytotoxic exposures causing regenerative proliferation or to mitogenic influences (Cohen and Ellwein, 1990). Increased cell division may elevate mutation by clonal expansion of initiated cells or by increasing the number of genetic errors by rapid cell division and reduced time for DNA repair. Some agents might result in an elevation of mutations by interfering with the enzymes involved in DNA repair and recombination (Barrett and Lee, 1992). Damage to certain critical DNA repair genes or other genes (e.g., the p53 gene) may result in genomic instability, which predisposes cells to further genetic alterations and increases the probability of neoplastic progression (Harris and Hollstein, 1993; Levine, 1994). Likewise, DNA repair processes may be saturated at certain doses of a

chemical, and thus result in an elevation of genetic alterations. Programmed cell death (apoptosis) can potentially be blocked by an agent, thereby permitting replication of cells carrying genetic errors. For example, peroxisome proliferators may act by suppressing apoptotic pathways (Shulte-Hermann et al., 1993; Bayly et al., 1994). At certain doses an agent may also generate reactive oxygen species that produce oxidative damage to DNA and other important macromolecules (Kehrer, 1993; Clayson et al., 1994; Chang et al., 1988). The role of these adducts, attributable to oxidative damage (e.g., 8-hydroxyguanine), in tumorigenesis is currently unclear.

Several carcinogens have been shown to induce aneuploidy (Gibson et al., 1995; Barrett, 1992). The loss or gain of chromosomes (i.e., aneuploidy) can result in the loss of heterozygosity or genomic instability (Fearon and Vogelstein, 1990; Cavenee et al., 1986). Agents that cause aneuploidy typically interfere with the normal process of chromosome segregation by interacting with non-DNA targets such as the proteins needed for chromosome movement. All tumors (except leukemias and lymphomas) are aneuploid, but whether this is the cause or the effect of tumorigenesis is not clear. Thus, it is important to understand whether the agent induces aneuploidy as a key early event in the carcinogenic process or is necessary for tumor progression.

It is possible for an agent to alter gene expression by transcriptional, translational, or post-translational modifications (Barrett, 1995). For example, perturbation of DNA methylation patterns may cause effects that contribute to carcinogenesis (Jones, 1986; Goodman and Counts, 1993; Holliday, 1987; Chuang et al., 1996). Overexpression of genes by DNA amplification has been observed in certain tumors (Vainio et al., 1992). Gene amplification may result from disproportionate DNA replication. Other mechanisms of altering gene expression may involve cellular reprogramming through hormonal or receptor-mediated mechanisms (Ashby et al., 1994; Barrett, 1992).

Both cell proliferation and programmed cell death are mandatory for the maintenance of homeostasis in normal tissue, and when altered become important elements of the carcinogenic process. The balance between the two directly affects the survival and growth of initiated cells, as well as preneoplastic and tumor cell populations (i.e., increase in cell proliferation or decrease in cell death) (Bellamy et al., 1995; Cohen and Ellwein, 1990, 1991; Cohen et al., 1991). Thus, measures of these events contribute to the weight of the evidence for cancer hazard and to mode-of-action understanding. In studies of proliferative effects, distinctions should be made between mitogenesis and regenerative proliferation (Cohen and Ellwein, 1990, 1991; Cohen et al., 1991). In applying information from studies on cell proliferation and apoptosis to risk assessment, it is important to identify the tissues and target cells involved, to measure effects in both normal and

neoplastic tissue, to distinguish between apoptosis and necrosis, and to determine the dose that affects these processes. Gap-junctional intercellular communication is believed to play a role in tissue and organ development and in the maintenance of a normal cellular phenotype within tissues. A growing body of evidence suggests that chemical interference with gap-junctional intercellular communication is a contributing factor in tumor development (Swierenga and Yamasaki, 1992; Yamasaki, 1995).

2.3.5.3. *Experimental Considerations in Evaluating Data on Precursor Events*

Most testing schemes for mutagenicity and other short-term assays were designed for hazard identification purposes; thus, these assays are generally conducted using acute exposures. For data on "precursor steps" to be useful in informing the dose-response curve for tumor induction below the level of observation, it is important that data come from in vivo studies where exposure is repeated or given over an extended period of time. Although consistency of results across different assays and animal models provides a stronger basis for drawing conclusions, it is desirable to have data on the precursor event in the same target organ, sex, animal strain, and species as the tumor data. In evaluating an agent's mode of action, it is usually not sufficient to determine that some event commences upon dosing. It is important to understand whether it is a causal event that plays a key role in the process that leads to tumor development, versus an effect of the cancer process itself or simply an associated event.

2.3.5.4. *Judging Data*

Criteria that are applicable for judging the adequacy of mechanistically based data include the following:

- mechanistic relevance of the data to carcinogenicity,
- number of studies of each endpoint,
- consistency of results in different test systems and different species,
- similar dose-response relationships for tumor and mode of action-related effects,
- tests conducted in accordance with generally accepted protocols, and
- degree of consensus and general acceptance among scientists regarding interpretation of the significance and specificity of the tests.

Although important information can be gained from in vitro test systems, a higher level of confidence is generally given to data that are derived from in vivo systems, particularly those

results that show a site concordance with the tumor data.

2.4. BIOMARKER INFORMATION

Various endpoints can serve as biological markers of events in biological systems or samples. In some cases, these molecular or cellular effects (e.g., DNA or protein adducts, mutation, chromosomal aberrations, levels of thyroid stimulating hormone) can be measured in blood, body fluids, cells, and tissues to serve as biomarkers of exposure in both animals and humans (Callemen et al., 1978; Birner et al., 1990). As such, they can do the following:

- act as an internal surrogate measure of chemical dose, representing as appropriate, either recent (e.g., serum concentration) or accumulated (e.g., hemoglobin adducts) exposure;
- help identify doses at which elements of the carcinogenic process are operating;
- aid in interspecies extrapolations when data are available from both experimental animal and human cells; and
- under certain circumstances, provide insights into the possible shape of the dose-response curve below levels where tumor incidences are observed (e.g., Choy, 1993).

Genetic and other findings (like changes in proto-oncogenes and tumor suppressor genes in preneoplastic and neoplastic tissue or, possibly, measures of endocrine disruption) can indicate the potential for disease and as such serve as biomarkers of effect. They, too, can be used in different ways:

- The spectrum of genetic changes in proliferative lesions and tumors following chemical administration to experimental animals can be determined and compared with those in spontaneous tumors in control animals, in animals exposed to other agents of varying structural and functional activities, and in persons exposed to the agent under study.
- They may provide a linkage to tumor response.
- They may help to identify subpopulations of individuals who may be at an elevated risk for cancer, e.g., cytochrome P450 2D6/debrisoquine sensitivity for lung cancer (Caporaso et al., 1989) or inherited colon cancer syndromes (Kinzler et al., 1991; Peltomäki et al., 1993).
- As with biomarkers of exposure, it may be justified in some cases to use these endpoints for dose-response assessment or to provide insight into the potential shape

of the dose-response curve at doses below those at which tumors are induced experimentally.

In applying biomarker data to cancer assessment (particularly assessments based on epidemiologic data), one should consider the following:

- routes of exposure,
- exposure to mixtures,
- time after exposure,
- sensitivity and specificity of biomarkers, and
- dose-response relationships.

2.5. MODE OF ACTION-GENERAL CONSIDERATIONS AND FRAMEWORK FOR ANALYSIS

2.5.1. General Considerations

The interaction of the biology of the organism and the chemical properties of the agent determine whether there is an adverse effect. Thus, mode-of-action analysis is based on physical, chemical, and biological information that helps to explain key events[2] in an agent's influence on development of tumors. The entire range of information developed in the assessment is reviewed to arrive at a reasoned judgment. An agent may work by more than one mode of action both at different sites and at the same tumor site. It is felt that at least some information bearing on mode of action (e.g., SAR, screening tests for mutagenicity) is present for most agents undergoing assessment of carcinogenicity, even though certainty about exact molecular mechanisms may be rare.

Inputs to mode-of-action analysis include tumor data in humans, animals, and among structural analogues as well as the other key data. The more complete the data package and generic knowledge about a given mode of action, the more confidence one has and the more one can replace or refine default science policy positions with relevant information. Making reasoned judgments is generally based on a data-rich source of chemical, chemical class, and tumor type-specific information. Many times there will be conflicting data and gaps in the information base; one must carefully evaluate these uncertainties before reaching any conclusion.

---

[2]A "key event" is an empirically observable, precursor step that is itself a necessary element of the mode of action, or is a marker for such an element.

In making decisions about potential modes of action and the relevance of animal tumor findings to humans (Ashby et al., 1990), very often the results of chronic animal studies may give important clues. Some of the important factors to review include the following:

- tumor types, e.g., those responsive to endocrine influence or those produced by reactive carcinogens (Ashby and Tennant, 1991);
- number of tumor sites, sexes, studies, and species affected or unaffected (Tennant, 1993);
- influence of route of exposure, spectrum of tumors, and local or systemic sites;
- target organ or system toxicity, e.g., urinary chemical changes associated with stone formation, effects on immune surveillance;
- presence of proliferative lesions, e.g., hepatic foci, hyperplasias;
- progression of lesions from preneoplastic to benign to malignant with dose and time;
- ratio of malignant to benign tumors as a function of dose and time;
- time of appearance of tumors after commencing exposure;
- tumors invading locally, metastasizing, producing death;
- tumors at sites in laboratory animals with high or low spontaneous historical incidence;
- biomarkers in tumor cells, both induced and spontaneous, e.g., DNA or protein adducts, mutation spectra, chromosome changes, oncogene activation; and
- shape of the dose response in the range of tumor observation, e.g., linear vs. profound change in slope.

Some of the myriad of ways that information from chronic animal studies influences mode-of-action judgments include the following. Multisite and multispecies tumor effects are often associated with mutagenic agents. Tumors restricted to one sex/species may suggest an influence restricted to gender, strain, or species. Late onset of tumors that are primarily benign or are at sites with a high historical background incidence or show reversal of lesions on cessation of exposure may point to a growth-promoting mode of action. The possibility that an agent may act differently in different tissues or have more than one mode of action in a single tissue must also be kept in mind.

Simple knowledge of sites of tumor increase in rodent studies can give preliminary clues as to mode of action. Experience at the National Toxicology Program (NTP) indicates that substances that are DNA reactive and produce gene mutations may be unique in producing tumors in certain anatomical sites, while tumors at other sites may arise from both mutagenic or

nonmutagenic influences (Ashby and Tennant, 1991; Huff et al., 1991).

Effects on tumor sites in rodents and other mode-of-action information has been explored for certain agents (Alison et al., 1994; Clayson, 1989; ECETOC, 1991; MacDonald et al., 1994; McClain, 1994; Tischler et al., 1991; ILSI, 1995; Cohen and Ellwein, 1991; FASEB, 1994; Havu et al., 1990; U.S. EPA, 1991c; Li et al., 1987; Grasso and Hinton, 1991; Larson et al., 1994; IARC, 1990; Jack et al., 1983; Stitzel et al., 1989; Ingram and Grasso, 1991; Bus and Popp, 1987; Prahalada et al., 1994; Yamada et al., 1994; Hill et al., 1989; Burek et al., 1988).

2.5.2. Evaluating a Postulated Mode of Action

Peer Review

This section contains a framework for evaluating a postulated mode of action. In reaching conclusions, the question of "general acceptance" of a mode of action will be tested as part of the independent peer review that EPA obtains for its assessment and conclusions. In some cases the mode of action may have already been established by development of a large body of research information and characterization of the phenomenon over time. In some cases there will have been development of an Agency policy, e.g., male rat thyroid disruption, or a series of previous assessments in which both the mode of action and its applicability to particular cases has been explored, e.g., urinary bladder stones. If so, the assessment and its peer review can be focused on the evidence that a particular agent acts in this mode.

In other cases, the mode of action previously may not have been the subject of an Agency document. If so, the data to support both the mode of action and the activity of the agent with respect to it will be the subjects of EPA assessment and subsequent peer review.

Use of the Framework

The framework supports a full analysis of mode-of-action information, but can also be used as a screen to decide whether sufficient information is available to evaluate or the data gaps are too substantial to justify further analysis. Mode-of-action conclusions are used to address the question of human relevance of animal tumor responses, to address differences in anticipated response among humans such as between children and adults or men and women, and as the basis of decisions about the anticipated shape of the dose-response relationship. Guidance on the latter appears in Section 3.

2.5.3. Framework for Evaluating a Postulated Carcinogenic Mode(s) of Action

This framework is intended to be an analytic tool for judging whether available data support a mode of carcinogenic action postulated for an agent. It is based upon considerations for causality in epidemiologic investigations originally articulated by Hill, but later modified by others and extended to experimental studies. The original Hill criteria were applied to epidemiologic data, while this framework is applied to a much wider assortment of experimental data, so it retains the basic principles of Hill but is much modified in content.

A mode of action is composed of key events and processes starting with the interaction of an agent with a cell, through operational and anatomical changes, resulting in cancer formation. "Mode" of action is contrasted with "mechanism" of action, which implies a more detailed, molecular description of events than is meant by mode of action. There are many examples of possible modes of carcinogenic action, such as mutagenicity, mitogenesis, inhibition of cell death, cytotoxicity with reparative cell proliferation, and immune suppression. All pertinent studies are reviewed in analyzing a mode of action, and an overall weighing of evidence is performed, laying out the strengths, weaknesses, and uncertainties of the case as well as potential alternative positions and rationales. Identifying data gaps and research needs is also part of the assessment.

To show that a postulated mode of action is operative, it is generally necessary to outline the sequence of events leading to cancer, to identify key events that can be measured, and to weigh information to determine whether there is a causal relationship between events and cancer formation. In no case will it be expected that the complete sequence is known at the molecular level. Instead, empirical observations made at different levels of biological organization are analyzed: biochemical, cellular, physiological, tissue, organ, and system levels.

Several important points should be kept in mind when working with the framework:

- The topics listed for analysis should *not* be regarded as a checklist of necessary "proofs." The judgment whether a postulated mode of action is supported by available data takes account of the analysis as a whole.
- The framework provides a structure for organizing the facts upon which conclusions as to mode of action rest. The purpose of using the framework is to make analysis transparent and allow the reader to understand the facts and reasoning behind a conclusion.
- The framework does not dictate an answer. The weight of evidence that is sufficient to support a decision about a mode of action may be less or more depending on the purpose of the analysis, e.g., screening, research needs identification, or full risk

assessment. To make the reasoning transparent, the purpose of the analysis ought to be made apparent to the reader.

- Toxicokinetic studies may contribute to mode-of-action analysis by identifying the active form of an agent that is central to the mode of action. Apart from contributing in this way, toxicokinetics studies may reveal effects of saturation of metabolic processes. These are not considered key events in a mode of action, but are given separate consideration in assessing dose metrics and potential nonlinearity of the dose-response relationship.
- Generally, "sufficient" support is a matter of scientific judgment in the context of the requirements of the decision maker or in the context of science policy guidance regarding a certain mode of action.
- While a postulated mode of action may be supported for a described response in a specific tissue, it may not explain other tumor responses observed. The latter will need separate consideration in hazard and dose-response assessment.

It is anticipated that in a risk assessment document, the analysis of a postulated mode of action will be presented before or with the characterization of an agent's potential hazard to humans.

2.5.3.1. ***Content of the Framework***

The framework analysis begins with a summary description of the postulated mode of action for a tumor type. (Each postulated mode of action requires separate analysis.) This is followed by topics for analysis and presentation in a convenient order. For illustration, the explanation of each topic includes typical questions to be addressed to the available empirical data and experimental observations anticipated to be pertinent. The latter will vary from case to case. For a particular mode of action, certain observations may be established as essential in practice or policy, e.g., measures of thyroid hormone levels in supporting thyroid hormone elevation as a key event in carcinogenesis. A conclusion and an analysis of human relevance including subpopulations are the the final parts of the analysis.

1. Summary Description of Postulated Mode of Action

This description briefly explains the sequence of events and processes that are considered to lead to cancer formation. For example, for thyroid disruption and thyroid follicular cell tumors:

Thyroid hormone production is regulated by actions of the hypothalamus,

pituitary, and thyroid gland. Homeostasis of thyroid hormone is maintained by a feedback loop between the hypothalamus and pituitary and the thyroid gland. The hypothalamus produces thyrotrophin reducing hormone (TRH), which stimulates the pituitary to produce thyroid stimulating hormone (TSH) which, in turn, stimulates the thyroid to produce thyroid hormone. The hypothalamus and pituitary respond to high levels of circulating thyroid hormone by suppressing TRH and TSH production, and to a low level by increasing them. The mode of action considered is continuous elevation of TSH levels that stimulates the thyroid gland to deplete its stores of thyroid hormone and continues to push production resulting in hypertrophy of the production cells (follicular cells) leading to hyperplasia, nodular hyperplasia, and, eventually, tumors of these cells. In rats, the chain of events may be induced by direct effects on hormone synthesis or by metabolic removal of circulating hormone.

2. "Identification of key events" is a consideration devised for this framework. A "key event" is an empirically observed precursor step consistent with a mode of action. In order to judge how well data support involvement of an event in carcinogenic processes, the experimental definition of the event or events must be clear and repeatable. To support an association, experiments need to define and measure an event consistently.

- Can a list of events be identified that are key to the carcinogenic process?
- Are the events well defined?

Pertinent observations: e.g., increased cell growth, organ weight, histology, proliferation assays, hormone or other protein perturbations, receptor-ligand changes, DNA or chromosome effects, cell cycle effects.

3. "Strength, consistency, specificity of association": A statistically significant association between events and a tumor response observed in well-conducted studies is supportive of causation. Consistent observations in a number of such studies with differing experimental designs increases that support, since different designs may reduce unknown biases. Studies showing "recovery," i.e., absence or reduction of carcinogenicity when the event is blocked or diminished, are particularly important tests of the association. Specificity of the association, without evidence of other modes of action, strengthens a causal conclusion.

- What is the level of statistical and biological significance for each event and for cancer?
- Do independent studies and different experimental hypothesis-testing approaches produce the same associations?
- Does the agent produce effects other than postulated?
- Is the key event associated with precursor lesions?

Pertinent observations: e.g., tumor response associated with events (site of action logically relates to event[s]), precursor lesions associated with events, initiation-promotion studies, stop/recovery studies.

4. "Dose-response relationship": If a key event and tumor endpoints increase with dose, a causal association can be strengthened. Dose-response associations of the key event with other precursor events can add further strength. Difficulty arises when an event is not causal, but accompanies the process generally. Dose-response studies coupled with mechanistic studies can assist in clarifying these relationships.

- What are the correlations among doses producing events and cancer?

Pertinent observations: e.g., 2-year bioassay observation of lesions correlated with observations of hormone changes and the same lesions in shorter term studies or in interim sacrifice.

5. "Temporal relationship": If an event is a cause of tumorigenesis, it must precede tumor appearance. An event may also be observed contemporaneously or after tumor appearance; these observations may add to the strength of association, but not to the temporal association.

- What is the ordering of events that underlie the carcinogenic process?
- Is this ordering consistent among independent studies?

Pertinent observations: Studies of varying duration observing the temporal sequence of events and tumorigenicity.

6. "Biological plausibility and coherence": The postulated mode of action and the

events that are part of it need to be based on current understanding of the biology of cancer to be accepted. If the body of information under scrutiny is consistent with other examples (including structurally related agents) for which the postulated mode of action is accepted, the case is strengthened. Since some modes of action can be anticipated to evoke effects other than cancer, the available toxicity database on noncancer effects can contribute to this evaluation, e.g., reproductive effects of certain hormonal disturbances.

- Is the mode of action consistent with what is known about carcinogenesis in general and for the case specifically?
- Are carcinogenic effects and events consistent across structural analogues?
- Is the database on the agent internally consistent in supporting the purported mode of action, including relevant noncancer toxicities?

Pertinent observations: Scientific basis for considering a postulated mode of action generally, given current state of knowledge of carcinogenic processes; previous examples of data sets showing the mode of action; data sets on analogues; coherence of data in this case from cancer and noncancer toxicity studies.

7. "Other modes of action": This discussion covers alternative modes of action for the tumor response considered and whether they are supported by the data. In addition, it provides a place to discuss other tumor observations that may be arising from a different mode of action than postulated.

8. "Conclusion": This is a brief conclusion and rationale as to whether the postulated mode of action is supported, also reflecting the purpose of the evaluation. The conclusion that a mode of action is supported is stronger as more of the above topic analyses point in the same direction, and weaker as fewer do so. The testing of the mode of action hypothesis by various experimental approaches with the same result creates a stronger basis for conclusions. Characteristics of strength of support include data showing that all key events are in sequence prior to tumor formation, dose and timing are consistent with the sequence, and reversal or reduction of key events and effects occurs upon cessation of dosing. The conclusion should address whether key event or associated metabolic information allows identification of rate-limiting measures of either the mode of action or of toxicokinetics.

9. "Human relevance, including subpopulations" : This is an analysis of data on the question whether a mode of action found to be operative in animals is also operative in humans and whether any human subpopulation is apt to qualitatively respond to the mode of action differently than the general population. Relevance to humans of animal responses is the default assumption since metazoans appear to share the basic modes of carcinogenic action.

When sufficient information is developed in mature animals to show a mode of action for a specific tumor type, an evaluation will be made of whether the mode of action is qualitatively applicable to children (including infants and fetuses), i.e., same sequence of key events is anticipated to be involved. Ideally we would have data pertinent to the question with respect to the agent under assessment. In the absence of such data, a cogent biological rationale needs to be developed regarding whether the mode of action is applicable to children. For the latter, the evaluation would cover the scientific information at large, including such considerations as age-related similarities and differences in the occurrence of the specific tumor type in the U.S. population, in occurrence of identified key events of the mode of action, in pertinent biochemical, physiological and toxicological processes, and in metabolism and excretion of the agent. Examples are given in case examples for chemicals T and Z in Appendix D. Based on the similarities of tumors following exposure to radiation, pharmaceuticals and viruses, from a qualitative standpoint, it may be anticipated that the same kind of tumors may develop following childhood or adult exposure to environmental chemicals. However. when there are no agent-specific data or there is not a cogent rationale supporting the comparability between responses in children and adults, the mode of action will not be considered to be applicable for children. It should also be noted that from a quantitative perspective, the same key events may lead to greater or lesser occurrence at different agents due to toxicokinetic and exposure considerations. These considerations need separate evaluation and may result in separate risk estimates for the young or for that portion of a lifetime.

2.6. WEIGHT-OF-EVIDENCE EVALUATION FOR POTENTIAL HUMAN CARCINOGENICITY

A weight-of-evidence evaluation is a collective evaluation of all pertinent information so that the full impact of biological plausibility and coherence is adequately considered. Identification and characterization of human carcinogenicity is based on human and experimental data, the nature, advantages, and limitations of which have been discussed in the preceding

sections.

The subsequent sections outline: (1) the basics of weighing individual lines of evidence and combining the entire body of evidence to make an informed judgment, and (2) classification descriptors of cancer hazard.

2.6.1. Weight-of-Evidence Analysis

Judgment about the weight of evidence involves considerations of the quality and adequacy of data and consistency of responses induced by the agent in question. The weight-of-evidence judgment requires combined input of relevant disciplines. Initial views of one kind of evidence may change significantly when other information is brought to the interpretation. For example, a positive animal carcinogenicity finding may be diminished by other key data; a weak association in epidemiologic studies may be bolstered by consideration of other key data and animal findings. Factors typically considered are illustrated in figures below. Generally, no single weighing factor on either side determines the overall weight. The factors are not scored mechanically by adding pluses and minuses; they are judged in combination.

Human Evidence

Analyzing the contribution of evidence from a body of human data requires examining available studies and weighing them in the context of well-accepted criteria for causation (see Section 2.2.1). A judgment is made about how closely the studies satisfy these criteria, individually and jointly, and how far they deviate from them. Existence of temporal relationships, consistent results in independent studies, strong association, reliable exposure data, presence of dose-related responses, freedom from biases and confounding factors, and high level of statistical significance are among the factors leading to increased confidence in a conclusion of causality.

Generally, the weight of human evidence increases with the number of adequate studies that show comparable results on populations exposed to the same agent under different conditions. The analysis takes into account all studies of high quality, whether showing positive associations or null results, or even protective effects. In weighing positive studies against null studies, possible reasons for inconsistent results should be sought, and results of studies that are judged to be of high quality are given more weight than those from studies judged to be methodologically less sound. See Figure 2-1.

Increase weight	Decrease weight
Number of independent studies with consistent results	Few studies
	Equally well-designed and conducted studies with null results
Most causal criteria satisfied:	Few causal criteria satisfied
Temporal relationship	
Strong association	
Reliable exposure data	
Dose-response relationship	
Freedom from bias and confounding	
Biological plausibility	
High statistical significance	

Figure 2-1. Factors for weighing human evidence.

Generally, no single factor is determinative. For example, strength of association is one of the causal criteria. A strong association (i.e., a relatively large risk) is more likely to indicate causality than a weak association. However, finding of a large excess risk in a single study must be balanced against the lack of consistency as reflected by null results from other equally well-designed and well-conducted studies. In this situation, the positive association of a single study may either suggest the presence of chance, bias, or confounding, or reflect different exposure conditions. On the other hand, evidence of weak but consistent associations across several studies suggests either causality or that the same confounder may be operating in all of these studies.

Animal Evidence

Evidence from long-term or other carcinogenicity studies in laboratory animals constitutes the second major class of information bearing on carcinogenicity. See Figure 2-2. As discussed in Section 2.2.2, each relevant study must be reviewed and evaluated as to its adequacy of design and conduct as well as the statistical significance and biological relevance of its findings. Factors that usually increase confidence in the predictivity of animal findings are those of (1) multiplicity of observations in independent studies; (2) severity of lesions, latency, and lesion progression; and (3) consistency in observations.

Increase weight	Decrease weight
Number of independent studies with consistent results	Single study
	Inconsistent results
Same site across species, structural analogues	
Multiple observations	Single site/species/sex
Species	
Sites	
Sexes	
Severity and progression of lesions	Benign tumors only
Early-in-life tumors/malignancy	
Dose-response relationships	High background of incidence tumors
Lesion progression	
Uncommon tumor	
Route of administration like human exposure	Route of administration unlike human exposure

Figure 2-2. Factors for weighing animal evidence.

Other Key Data

Additional information bearing on the qualitative assessment of carcinogenic potential may be gained from comparative pharmacokinetic and metabolism studies, genetic toxicity studies, SAR analysis, and other studies of an agent's properties. See Figure 2-3. Information from these studies helps to elucidate potential modes of action and biological fate and disposition. The knowledge gained supports interpretation of cancer studies in humans and animals and provides a separate source of information about carcinogenic potential.

Increase weight	Decrease weight
A rich set of other key data are available	Few or poor data
Physicochemical data	or
Data indicate reactivity with macromolecules	Inadequate data necessitate use of default assumptions
Structure-activity relationships support hazard potential	or
Comparable metabolism and toxicokinetics between species	Data show that animal findings are not relevant to humans
Toxicological and human clinical data support tumor findings	
Biomarker data support attribution of effects to agent	
Mode-of-action data support causal interpretation of human evidence or relevance of animal evidence	

Figure 2-3. Factors for weighing other data.

Totality of Evidence

In reaching a view of the entire weight of evidence, all data and inferences are merged. Figure 2-4 indicates the generalities. In fact, possible weights of evidence span a broad continuum that cannot be capsulized. Most of the time the data in various lines of evidence fall in the middle of the weights represented in the four figures in this section.

Increase Weight	Decrease Weight
Evidence of human causality	Data not available or do not show causality
Evidence of animal effects relevant to humans	Data not available or not relevant
Coherent inferences	Conflicting data
Comparable metabolism and toxicokinetics between species	Metabolism and toxicokinetics not comparable
Mode of action comparable across species	Mode of action not comparable across species

Figure 2-4. Factors for weighing totality of evidence.

The following section and the weight-of-evidence narrative discussed in Section 2.8 provide a way to state a conclusion and capture this complexity in a consistent way.

2.6.2. Descriptors for Summarizing Weight of Evidence

To express conclusions about the weight of evidence for human carcinogenic potential, standard descriptors are utilized as part of the narrative (see Section 2.7.2.). The descriptors are not meant to replace an explanation of the nuances of the biological evidence, but rather to summarize it. Applying a descriptor is a matter of judgment and cannot be reduced to a formula. Each standard descriptor may be applicable to a wide variety of potential data sets and weights of evidence. There will always be gray areas, gradations, and borderline cases. That is why the descriptors are presented only in the context of a weight of evidence narrative. Using them within a narrative preserves and presents the complexity that is an essential part of the hazard characterization. Risk managers should consider the entire range of information included in the narrative rather than focusing simply on the descriptor.

Different conclusions may be reached for a single agent when carcinogenicity is dose or route dependent. For instance, the agent is likely to be carcinogenic by one route of exposure but not by others (Section 2.3.3). In this instance, more than one descriptor is used, one for each route of exposure. Another example would be that an agent is likely carcinogenic above a certain dose range but not likely to be carcinogenic below that range.

The descriptors are standardized and are to be used consistently from case to case. They are part of the first sentence of the narrative. The discussions below explain descriptors which appear in italics, and along with Appendices A and C, illustrate their use, including by route of exposure.

"*Carcinogenic To Humans*"

This descriptor is appropriate when there is convincing epidemiologic evidence demonstrating causality between human exposure and cancer.

This descriptor is also appropriate when there is an absence of conclusive epidemiologic evidence to clearly establish a cause and effect relationship between human exposure and cancer, but there is compelling evidence of carcinogenicity in animals and mechanistic information in animals and humans demonstrating similar mode(s) of carcinogenic action. It is used when all of the following conditions are met:

- There is evidence in a human population(s) of association of exposure to the agent with cancer, but not enough to show a causal association, and
- There is extensive evidence of carcinogenicity, and
- The mode(s) of carcinogenic action and associated key events have been identified in animals, and
- The keys events that precede the cancer response in animals have been observed in the human population(s) that also shows evidence of an association of exposure to the agent with cancer.

"Likely To be Carcinogenic To Humans"

This descriptor is appropriate when the available tumor effects and other key data are adequate to demonstrate carcinogenic potential to humans. Adequate data are within a spectrum. At one end is evidence for an association between human exposure to the agent and cancer and strong experimental evidence of carcinogenicity in animals; at the other, with no human data, the weight of experimental evidence shows animal carcinogenicity by a mode or modes of action that are relevant or assumed to be relevant to humans.

"Suggestive Evidence of Carcinogenicity,
but Not Sufficient to Assess Human Carcinogenic Potential"

This descriptor is appropriate when the evidence from human or animal data is suggestive of carcinogenicity, which raises a concern for carcinogenic effects but is judged not sufficient for a conclusion as to human carcinogenic potential. Examples of such evidence may include: a marginal increase in tumors that may be exposure-related, or evidence is observed only in a single study, or the only evidence is limited to certain high background tumors in one sex of one species. Dose-response assessment is not indicated for these agents. Further studies would be needed to determine human carcinogenic potential.

"Data Are Inadequate for An Assessment of Human Carcinogenic Potential"

This descriptor is used when available data are judged inadequate to perform an assessment. This includes a case when there is a lack of pertinent or useful data or when existing evidence is conflicting, e.g., some evidence is suggestive of carcinogenic effects, but other equally

pertinent evidence does not confirm a concern.

"Not likely To Be Carcinogenic To Humans"

This descriptor is used when the available data are considered robust for deciding that there is no basis for human hazard concern. The judgment may be based on—

- Extensive human experience that demonstrates lack of carcinogenic effect (e.g., phenobarbital).

- Animal evidence that demonstrates lack of carcinogenic effect in at least two well-designed and well-conducted studies in two appropriate animal species (in the absence of human data suggesting a potential for cancer effects).

- Extensive experimental evidence showing that the only carcinogenic effects observed in animals are not considered relevant to humans (e.g., showing only effects in the male rat kidney due to accumulation of α2u-globulin).

- Evidence that carcinogenic effects are not likely by a particular route of exposure (Section 2.3.3.)
- Evidence that carcinogenic effects are not anticipated below a defined dose range.

2.7. TECHNICAL HAZARD CHARACTERIZATION

The hazard characterization has two functions. First, it presents results of the hazard assessment and an explanation of how the weight-of-evidence conclusion was reached. It explains the potential for human hazard, anticipated attributes of its expression, and mode-of-action considerations for dose response. Second, it contains the information needed for eventual incorporation into a risk characterization consistent with EPA guidance on risk characterization (U.S. EPA, 1995).

The characterization summarizes the conclusions reached concerning the mode of action of the agent and devotes particular attention to a clear statement of the strengths and weaknesses of the inferences made and their relation to the framework for analyzing described in Chapter 2. The implications of the mode of action for the dose-response assessment are clearly stated, along

with the degree of confidence in those conclusions.

The characterization qualitatively describes the conditions under which the agent's effects may be expressed in human beings. These qualitative hazard conditions are ones that are observable in the tumor and other key data without having done either quantitative dose-response or exposure assessment. The description includes how expression is affected by route of exposure and dose levels and durations of exposure. Implications for disproportionate risks in particular subpopulations, including fetuses and children, are identified when such information exists.

The discussion of limitations of dose as a qualitative aspect of hazard addresses the question of whether reaching a certain dose range appears to be a precondition for a hazard to be expressed; for example, when carcinogenic effects are secondary to another toxic effect that appears only when a certain dose level is reached. The assumption is made that an agent that causes internal tumors by one route of exposure will be carcinogenic by another route, if it is absorbed by the second route to give an internal dose. Conversely, if there is a route of exposure by which the agent is not absorbed (does not cross an absorption barrier; e.g., the exchange boundaries of skin, lung, and digestive tract through uptake processes) to any significant degree, hazard is not anticipated by that route. An exception to the latter statement would be when the site of contact is also the target tissue of carcinogenicity. Duration of exposure may be a precondition for hazard if, for example, the mode of action requires cytotoxicity or a physiologic change, or is mitogenicity, for which exposure must be sustained for a period of time before effects occur. The characterization could note that one would not anticipate a hazard from isolated, acute exposures. The above conditions are qualitative ones regarding preconditions for effects, not issues of relative absorption or potency at different dose levels. The latter are dealt with under dose-response assessment (Section 3), and their implications can only be assessed after human exposure data are applied in the characterization of risk.

The characterization describes conclusions about mode-of-action information and its support for recommending dose-response approaches.

The hazard characterization routinely includes the following in support of risk characterization:

- a summary of results of the assessment;
- identification of the kinds of data available to support conclusions and explanation of how the data fit together, highlighting the quality of the data in each line of evidence, e.g., tumor effects, short-term studies, structure-activity relationships), and

highlighting the coherence of inferences from the different kinds of data;

- strengths and limitations (uncertainties) of the data and assessment, including identification of default assumptions invoked in the face of missing or inadequate data;
- identification of alternative interpretations of data that are considered equally plausible;
- identification of any subpopulations believed to be more susceptible to the hazard than the general population, especially attending to fetuses, infants, and children;
- conclusions about the agent's mode of action and recommended dose-response approaches; and
- significant issues regarding interpretation of data that arose in the assessment. Typical ones may include:
  - -- determining causality in human studies,
  - -- dosing (MTD), background tumor rates, relevance of animal tumors to humans;
  - -- weighing studies with positive and null results, considering the influence of other available kinds of evidence; and
  - -- drawing conclusions based on mode-of-action data versus using a default assumption about the mode of action.

2.8. WEIGHT-OF-EVIDENCE NARRATIVE

The weight-of-evidence narrative summarizes the results of hazard assessment employing the descriptors defined in Section 2.6.1. The narrative (about two pages in length) explains an agent's human carcinogenic potential and the conditions of its expression. If data do not allow a conclusion as to carcinogenicity, the narrative explains the basis of this determination. An example narrative appears below. More examples appear in Appendix A.

The items regularly included in a narrative are:

- name of agent and Chemical Abstracts Services number, if available;
- conclusions (by route of exposure) about human carcinogenicity, using a standard descriptor from Section 2.6.1;
- summary of human and animal tumor data on the agent or its structural analogues, their relevance, and biological plausibility;
- other key data (e.g., structure-activity data, toxicokinetics and metabolism, short-term studies, other relevant toxicity or clinical data);

- discussion of possible mode(s) of action and appropriate dose-response approach(es); and
- conditions of expression of carcinogenicity, including route, duration, and magnitude of exposure.

Example Narrative

Aromatic Compound

CAS# XXX

CANCER HAZARD SUMMARY

Aromatic compound (AR) is *carcinogenic to humans* by all routes of exposure.

The weight of evidence of human carcinogenicity is based on (a) consistent evidence of elevated leukemia incidence in studies of exposed workers and significant increases of genetic damage in bone marrow cells and blood lymphocytes of exposed workers; (b) significantly increased incidence of cancer in both sexes of several strains of rats and mice; (c) genetic damage in bone marrow cells of exposed rodents and effects on intracellular signals that control cell growth.

AR is readily absorbed by all routes of exposure and rapidly distributed throughout the body. The mode of action of AR is not understood. A dose-response assessment that assumes linearity of the relationship is recommended as a default.

SUPPORTING INFORMATION

Data include numerous human epidemiologic and biomonitoring studies, long-term bioassays, and other data on effects of AR on genetic material and cell growth processes. The key epidemiologic studies and animal studies are well conducted and reliable. The other data are generally of good quality also.

Human Effects

Numerous epidemiologic and case studies have reported an increased incidence or a causal relationship associating exposure to AR and leukemia. Among the studies are five for which the design and performance as well as follow-up are considered adequate to demonstrate the causal relationship. Biomonitoring studies of exposed workers have found dose-related increases in chromosomal aberrations in bone marrow cells and blood lymphocytes.

Animal Effects

AR caused increased incidence of tumors in various tissues in both sexes of several rat and mouse strains. AR also caused chromosomal aberrations in rabbits, mice, and rats--as it does in humans.

Other Key Data

AR itself is not DNA reactive and is not mutagenic in an array of test systems both in vitro and in vivo. Metabolism of AR yields several metabolites that have been separately studied for effects on carcinogenic processes. Some have mutagenic activity in test systems and some have other effects on growth controls inside cells.

MODE OF ACTION

No rodent tumor precisely matches human leukemia in pathology. The closest parallel is a mouse cancer of blood-forming tissue. Studies of the effects of AR at the cell level in this model system are ongoing. As yet, the mode of action of AR is unclear, but most likely the carcinogenic activity is associated with one or a combination of its metabolites. It is appropriate to apply a linear approach to the dose-response assessment pending a better understanding because: (a) genetic damage is a typical effect of AR exposure in mammals, and (b) metabolites of AR produce mutagenic effects in addition to their other effects on cell growth controls; AR is a multitissue carcinogen in mammals, suggesting that it is affecting a common controlling mechanism of cell growth.

3. DOSE-RESPONSE ASSESSMENT

Dose-response assessment evaluates potential risks to humans at exposure levels of interest. The approach to dose-response assessment for a particular agent is based on the conclusion reached as to its mode of action (Sections. 2.4 -2.5). The evaluation first covers the relationship of the dose[1] to the degree of response in the dose range of observation in experiments or human studies. This evaluation is then followed by extrapolation to estimate response at lower environmental exposure levels (ILSI, 1995). In general, three extrapolations may be made: from high to low doses, from animal to human responses, and from one route of exposure to another.

Cancer is a disease that develops through many cell and tissue changes over time. Traditional dose-response assessment procedures using tumor incidence as the response have seldom taken into account the effects of key events within the whole biological process, even though these events are the determinants of the overall dose-response. This has been due to lack of empirical data and understanding about these events. As more data become available and our understanding about how agents cause cancer improves, they can be used in dose-response assessment along with the traditional procedures. These guidelines encourage use of these new data as they become available to improve dose-response assessment.

In this discussion, "response" data include measures of key events[2] considered integral to the carcinogenic process, in addition to tumor incidence. These responses may include changes in DNA, chromosomes, or other key macromolecules; effects on growth signal transduction, including induction of hormonal changes; or physiological or toxic effects that affect cell proliferation. Key events are precursors to cancer pathology; they may include proliferative events diagnosed as precancerous, but not pathology that is judged to be cancer. Analysis of such responses may be done along with those of tumor incidence to enhance the tumor dose-response analysis. If dose-response analysis of non tumor key events is more informative about the carcinogenic process for an agent, it is used in lieu of, or in conjunction with, tumor incidence

---

1. For this discussion, "exposure" means contact of an agent with the outer boundary of an organism. "Applied dose" means the amount of an agent presented to an absorption barrier and available for absorption. "Internal dose" means the amount crossing an absorption barrier (e.g., the exchange boundaries of skin, lung, and digestive tract) through uptake processes. "Delivered dose" for an organ or cell means the amount available for interaction with that organ or cell (U.S. EPA, 1992a).

2. A "key event" is an empirically observed precursor consistent with a mode of action.

analysis for the overall dose-response assessment.

"Dose" means the "human equivalent dose" as discussed in Section 3.3, unless otherwise noted. When animal responses are used in the assessment, the animal dose is adjusted to human equivalence. The preferred approach for this is to use toxicokinetic modeling to compare species. If this is not possible given the data available, a default factor for allometric scaling of oral dose is provided. For adjustment of inhalation dose, the EPA's Reference Concentration (RfC) methodology is used.

Coverage of the Chapter

This chapter covers: 1) consideration of mode of action in selecting dose-response assessment approaches, 2) assessment of observed data and extrapolation procedures, 3) analyses of response data and 4) analyses of dose data. The final section discusses dose-response characterization.

3.1 HUMAN STUDIES

Analysis of human studies in the observed range is determined according to the type of study and how dose and response are measured in the study. In some cases the agent may have discernible interactive effects with another agent (e.g., asbestos and smoking), making possible estimation of contribution of the agent and others as risk factors. Also, in some cases, estimation of population risk in addition to, or in lieu of, individual risk may be appropriate. The following discussions are addressed mainly to animal data. Nevertheless, if human data permit, the principles or approaches below apply for performing dose-response assessment in two parts--range of observation and range of extrapolation, for deriving a point of departure, and for linear or margin of exposure analysis according to mode of action (NRC, 1999; Teta, 1999). The approach is tailored to the nature of the human data and the mode of action data available, if any.

3.2. MODE OF ACTION AND DOSE-RESPONSE APPROACH

The cancer dose-response relationship(s) for a chemical is considered in a two step process. First is the determination of the mode of action and dose response for each tumor type that results in a significant increase in tumor incidence. Second is an analysis of the information bearing on all tumor types that are increased in incidence by the chemical. The overall synthesis

includes consideration of the number of sites, their consistency across sexes, strains and species, the strength of the mode of action information for each tumor type, the anticipated relevance of each tumor type to humans, and the consistency of the means of estimating risks across tumor types.

For each tumor the mode of action and other information may support one of the following dose response extrapolations: 1) linear, 2) nonlinear using a margin of exposure (MOE) analysis, or 3) both linear and nonlinear (MOE) analyses. In rare cases, detailed mode of action information may be available which allow the formulation of a biologically based model. Examples include the following:

Factors Supporting a Linear Approach

Any of the following conclusions leads to selection of a linear dose-response assessment approach:

- There is an absence of sufficient tumor mode of action information.
- The chemical has direct DNA mutagenic activity or other indications of DNA effects that are consistent with linearity.
- Human exposure or body burden is high and near doses associated with key events in the carcinogenic process (e.g., 2,3,7,8-tetrachlorodibenzo-p-dioxin)
- Mode of action analysis does not support direct DNA effects, but the dose-response relationship is expected to be linear (e.g., certain receptor-mediated effects)

Factors Supporting a Nonlinear Approach

Any of the following conclusions leads to selection of a nonlinear (margin of exposure) approach to dose-response assessment:

- A tumor mode of action supporting nonlinearity applies (e.g., some cytotoxic and hormonal agents such as disruptors of hormone homeostasis), *and* the chemical does not demonstrate mutagenic effects consistent with linearity.
- A mode of action supporting nonlinearity has been demonstrated, *and* the chemical has some indication of mutagenic activity, but it is judged not to play a significant role in tumor causation.

Factors Supporting Both Linear And Nonlinear Approaches

Any of the following conclusions leads to selection of both a linear and nonlinear approach

to dose-response assessment. Relative support for each dose response method and advice on the use of that information needs to be presented. In some cases, evidence for one mode of action is stronger that for the other, allowing emphasis to be placed on that dose-response approach. In other cases, both modes of action are equally possible, and both dose-response approaches should be emphasized.

- Modes of action for a single tumor type support both linear and nonlinear dose response in different parts of the dose-response (e.g., 4,4' methylene chloride).
- A tumor mode of action supports different approaches at high and low dose; e.g., at high dose, nonlinearity, but, at low dose, linearity (e.g., formaldehyde).
- The agent is not DNA-reactive and all plausible modes of action are consistent with nonlinearity, but not fully established (arsenic).
- Modes of action for different tumor types support differing approaches, e.g., nonlinear for one and linear due to lack of mode of action for the other (e.g., trichloroethylene).

The use of biologically based models is covered below.

3.3. DOSE-RESPONSE ANALYSIS

3.3.1. Modeling the Overall Process--Biologically-based Models

Generally applicable biologically-based models may be applied such as the two-stage models of initiation plus clonal expansion and progression developed by Moolgavkar and Knudson (1981), Chen and Farland (1991) and others. These models of the carcinogenic process continue to be improved, but are not yet standard methods. No model of this kind is available for standard application.

If data are extensive and sufficient to quantitatively relate specific key events in the cancer process to neoplasia, and the purpose of the assessment is such as to justify investing the necessary resources, a biologically-based model may be developed on an agent-specific basis. Before developing such a model, extensive data are needed to build its form as well as to estimate how well it conforms with the observed data to support confidence in results. Theoretical estimates of critical parameters, such as cell proliferation rates, are not used to enable application of such a model in the absence of data (Portier, 1987). It is possible that different models will provide equivalent fits to the observed data but differ substantially in their projections below the observed range. This is often the case when a model is over-parameterized (that is, there are

more parameters to be estimated than data points to be fitted), so that different combinations of parameter estimates can yield similar results in the observed range. For this reason, critical parameters of a biologically based model, such as mutation and proliferation rates, are measured in the laboratory and not estimated by curve-fitting to tumor incidence data. This approach helps reduce model uncertainty (i.e., uncertainty due to choice of models or model structure) and ensures that the models do not give answers that are biologically unrealistic. This approach also provides a robustness of results (i.e., results are not likely to change substantially when fitted to a slightly different data set), if the mode of action is sufficiently understood so that model parameters represent rates and other quantities associated with known key events in tumor development.

Such models are to be distinguished from toxicokinetic models (i.e., physiologically based pharmacokinetic" models) which address dose issues, as discussed in Section 3.3.2. Effects on dose such as saturation of metabolic pathways may introduce nonlinearities in the dose-response relationship, but are not modes of action, and are dealt within arriving at an appropriate dose metric.

3.3.2. Analysis in the Range of Observation

This section covers use of information about key events which may be in the context of either human or animal data. It then discusses curve-fitting and selecting a point of departure with regard to animal data. Last, it discusses human data.

3.3.2.1. Applying Information About Key Events

Even though a biologically-based model may not be feasible, information about key events in the process can be used in the assessment. The principle underlying these Guidelines is to use approaches that include as much information about these events as possible. When such information is available, it may be used in a variety of ways:

1) If an event(s) is quantitatively described and considered key to cancer development, its dose-response assessment in the range of observation can be used in conjunction with, or in lieu of, the dose-response for tumor incidence to establish the point of departure for extrapolation. [Caution must be used in using rates of molecular events such as mutation or cell proliferation or of signal transduction. Such rates may be difficult to relate to cell or tissue changes overall. The timing of observations of these phenomena, as well as the cell type involved, need to be linked to other precursor events to ensure the measurement is truly a "key"event (see Section 2.5). In many cases such rates are more appropriately used as in "2)" or "3)" below.]

2) Quantitative description of a key event(s) can be used to test whether the dose-response for tumor incidence can be confidently extended to support a lower point of departure for linear extrapolation than the tumor data alone would support (e.g., to an LED_{01} from an LED_{10}).

3) Quantitative information on a key event(s) can be used to address the question of how quickly risk decreases as dose decreases in a margin of exposure analysis.

3.3.2.2. Procedures for Analysis in the Range of Observation of Animal Studies

Curve-fitting

A curve-fitting procedure is used that is appropriate to the kind of response data in the range of observation. This may be tumor incidence or data on a key event(s). For incidence information, the Agency applies a standard curve-fitting procedure to provide consistency among assessments. This procedure models incidence, adjusted for background, as an increasing function of dose; it is available to the public on the Agency's World Wide Web site for immediate use or for downloading (reference to be provided). The procedure identifies situations in which the standard algorithm fails to yield a reliable point of departure, signaling the need for additional judgment and an alternative analysis.

For tumor incidence studies that provide time-to-tumor information, more elaborate models would be appropriate. The Agency intends to provide a time-to-tumor version of its standard procedure in the future.

For non tumor data, curve-fitting procedures are used that are appropriate to the kind of response data in the observed range, and are explained in each case (reference to benchmark models to be provided).

NOAEL/LOAEL

As discussed below, the observed range of data may be represented by a NOAEL/LOAEL procedure when a margin of exposure analysis is chosen as the default procedure for nonlinear dose-response extrapolation.

3.3.2.3. Point of Departure for Extrapolation from Observed Animal Data

A point of departure from observed data--for tumor incidence, or for key event(s)--is estimated to mark the beginning of extrapolation. This is a point that is either a data point or an estimated point that can be considered to be in the range of observation, without significant extrapolation. Depending on the kind of data available and the purpose of the analysis, there are

differing procedures for estimating the point of departure. The point of departure employs the human equivalent dose.

Incidence data are most amenable to curve-fitting procedures. For example, tumor data from a rodent bioassay are traditionally modeled with curve-fitting procedures. Some key event data may also be in the form of incidence data (e.g., hyperplasia), but more likely will be continuous data for which currently there are not standard and consistent modeling procedures. Continuous data include, for instance, tissue weight changes or blood levels of a hormone. NOAEL/LOAEL procedures are available for continuous and other data as needed.

Point of Departure Using Data Suitable for Curve-fitting

When a curve-fitting procedure is applied to tumor data (see Figure 3-1) or to incidence data on a key event, the point of departure used in most cases is the LED_{10}--the 95% lower confidence limit on a dose associated with 10% extra risk adjusted for background. For tumor data, it is used as a matter of science policy to provide consistency among assessments. It is also useful in comparing results with assessment of noncancer endpoints (U.S. EPA, 1991d). The 10% level is selected because a 10% response is at or just below the limit of sensitivity for discerning a statistically significant tumor increase in most long-term rodent studies (Haseman, 1983), and is within the observed range for many other kinds of toxicity studies. Use of the lower limit takes experimental variability and sample size into account. If a tumor incidence study has greater than usual sensitivity and an observed response is below LED_{10}, then a lower point for linear extrapolation can be used to improve the assessment. [The ED_{10} (central estimate) is appropriate for use in relative hazard/potency ranking among agents for priority setting because it is a more confident comparison point among many assessments than an extrapolated point. Because of its convenience for comparison uses, the ED_{10} is always presented for reference with its upper and lower 95% confidence limits.]

The LED_{10} is adopted as the standard point of departure for non tumor key event or toxicity incidence data in order to harmonize curve-fitting procedures between cancer and noncancer toxicity assessments. Because the NOAEL in study protocols for non tumor toxicity can range from about a 5% to a 30% effect level (Faustman et al., 1994), adopting the 10% effect level as the standard point of departure will accommodate most of these data sets without departing the range of observation. The LED_{10} can be regarded as an improved and harmonized estimate of the NOAEL.

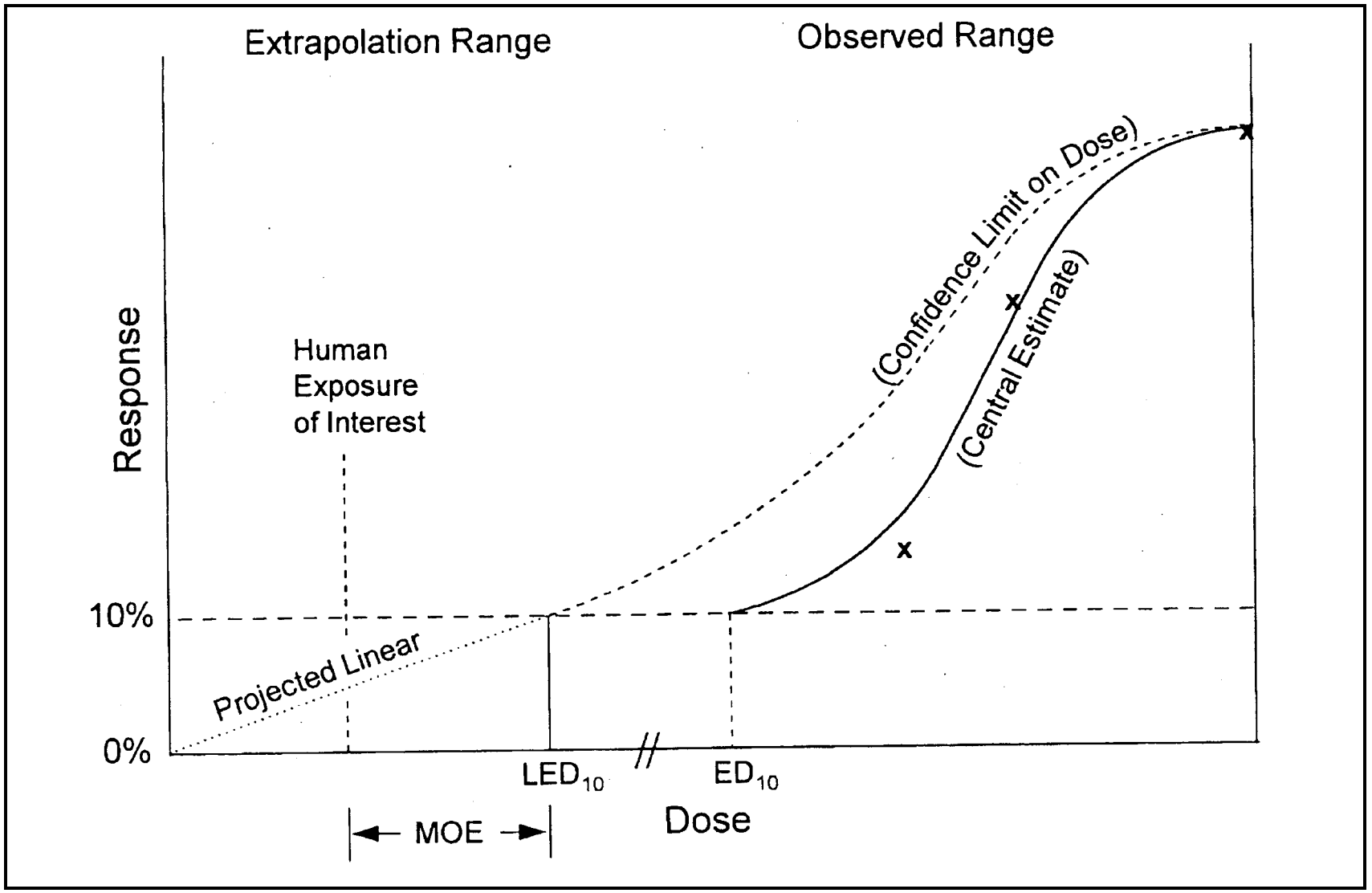

Figure 3-1. Graphical presentation of data and extrapolation.

Point of Departure Using Data Suitable for a NOAEL/LOAEL Procedure

The point of departure may be a NOAEL when a margin of exposure analysis is the nonlinear dose-response approach. The kinds of data available and the circumstances of the assessment both contribute to deciding to estimate a NOAEL or LOAEL which is not as rigorous or as ideal as curve-fitting, but can be appropriate. The NOAEL/LOAEL procedure is used to maintain consistency among assessments while still encouraging quantitative analyses of the data by modeling to explore underlying phenomena.

The circumstances of an assessment can also lead to choosing a NOAEL/LOAEL approach. If several data sets for key events and tumor response are available for an agent, and they are a mixture of continuous and incidence data, the most practicable way to assess them together is through a NOAEL/LOAEL approach. The purpose of the assessment also may lead to a decision to use the NOAEL/LOAEL approach. A preliminary or screening assessment to decide whether risk concern is high or low or to decide on additional data requirements is one example. Similarly, the nature of the regulatory decision may be served well by this approach to assessment.

3.3.3. Analysis in the Range of Extrapolation--Default Procedures

Extrapolation from the point of departure to lower doses is usually necessary, and in the absence of a data set rich enough to support a biologically based model, is conducted using one of the two default procedures described below. The Agency has adopted these procedures as a matter of science policy based on current hypotheses of the potential shapes of dose-response curves for differing modes of action at low doses. The choice of the procedure to be used in an individual case is a judgment based on the agent's mode of action (See Section 3.2).

3.3.3.1. Linear Procedure

For linear extrapolation, a straight line is drawn from the point of departure expressed as a human equivalent dose (Section 3.3.2) to the origin--zero incremental dose, zero incremental response to give a probability of extra risk. The slope of the line expresses extra risk per dose unit (Flamm and Winbush, 1984; Gaylor and Kodell, 1980; Krewski et al., 1984). Risk is the product of the slope and anticipated exposure. This approach to assessing risk is considered generally conservative of public health, including sensitive subpopulations, in the absence of specific information about the extent of human variability in sensitivity to effects. When a linear extrapolation procedure is used, the risk characterization summary also displays the degree of extrapolation from empirical data by showing the margin of exposure associated with exposure scenarios of interest as below.

3.3.3.2. Nonlinear Extrapolation

A default assumption of nonlinearity is appropriate when there is no evidence for linearity and sufficient evidence to support an assumption of nonlinearity. The mode of action may lead to a dose-response relationship that is nonlinear, with response falling much more quickly than linearly with dose, or being most influenced by individual differences in sensitivity. Alternatively, the mode of action may theoretically have a threshold, e.g., the carcinogenicity may be a secondary effect of toxicity or of an induced physiological change that is itself a threshold phenomenon (see Appendix C, example 5, or Appendix D, example 2). The EPA does not generally try to distinguish between modes of action that might imply a "true threshold" from others with a nonlinear dose-response relationship. Except in unusual cases where extensive information is available, it is not possible to distinguish between these empirically.

As a matter of science policy under this analysis, nonlinear probability functions are not fitted to the response data to extrapolate quantitative low-dose risk estimates because different models can lead to a very wide range of results, and there is currently no basis, generally, to choose among them. Thus, the default procedure for nonlinear extrapolation is to conduct a margin of exposure analysis, as described below, to evaluate concern for levels of exposure.

3.3.3.2.1. Margin of Exposure Analysis

A margin of exposure is defined as the point of departure divided by the environmental exposure of interest. The environmental exposures of interest, for which margins of exposure are estimated, may be actual or projected exposure levels. A risk manager decides whether a given margin of exposure is acceptable under applicable management policy criteria. The risk assessment provides supporting information to assist the decisionmaker in this determination.

A margin of exposure analysis is applicable if data are sufficient to presume a non-linear dose-response function containing a significant change in slope. If, in a particular case, the evidence indicates a biological threshold, as in the case of carcinogenicity being secondary to another toxicity that has a threshold, an RfD[3] or RfC like approach may be estimated and considered in cancer assessment. In this case, the RfD or RfC is an estimate with uncertainty

---

3. A reference dose (RfD) or reference concentration (RfC) for noncancer toxicity is an estimate with uncertainty spanning perhaps an order of magnitude of daily exposure to the human population (including sensitive subgroups) that is anticipated to be without appreciable deleterious effects during a lifetime. It is arrived at by dividing empirical data on effects by uncertainty factors that consider inter- and intraspecies variability, extent of data on all important chronic exposure toxicity endpoints, and availability of chronic as opposed to subchronic data.

spanning perhaps an order of magnitude of daily exposure to the human population (including sensitive subgroups) that is anticipated to be without a cancer hazard despite a lifetime of exposure. In many cases, data may be insufficient to determine an RfD and/or an RfC for the cancer endpoint. In that case, a margin of exposure analysis provides useful input to the decision-maker regarding the distance between an exposure of interest and the range of observation where cancer risk is inferred to be sub-linear.

To support a risk manager's consideration of the margin of exposure, all of the pertinent hazard, dose-response, and human exposure information is characterized so as to provide insights about the scientific community's current understanding of the phenomena that may be occurring as dose (exposure) decreases substantially below the observed data. The goal is to provide as much information as possible about the risk reduction that accompanies lowering of exposure and the adequacy of a margin of exposure based on scientific input, recognizing that, in some cases, legislative, sociological, and/or technological issues may also impact on the decision regarding the acceptability of a given margin of exposure. The discussion below describes the general principles and major elements to be considered in a margin of exposure analysis. The Agency will develop more specific guidance on the margin of exposure approach, as recommended (SAB, 1999). The guidance will be peer reviewed and published separately as part of the Agency's implementation activity of these guidelines.

For a margin of exposure analysis, the point of departure would ideally be the dose where the key events in tumor development would not occur in a heterogenous human population, thus representing an actual "no effect level." Therefore, it is recommended that margin of exposure analyses be based on precursor responses rather than tumor incidences, since precursor events can often be detected with greater sensitivity(i.e. both earlier and at lower doses), providing further input to the decision regarding acceptability of the margin of exposure. An analysis of an actual point of departure derived from available data, however, would often contain residual uncertainty regarding its designation as an actual no effect level for cancer in the population. The earlier the precursor event in the carcinogenic process and the larger the margin of exposure the more likely the exposure of interest will be without appreciable risk of cancer. To this end, some important points to address in the analysis of the point of departure and the margin of exposure include the following:

- *Nature of the response*. Is the point of departure based on tumors or on a *key event* that is

a precursor to tumors? A mode of action can be represented by a sequence of dose-response curves, where an early key event arises at a low dose, subsequent key events at higher doses, and tumors at a still higher dose. For example, a mode of action that begins with bladder stones and progresses through epithelial irritation and hyperplasia before producing tumors can be represented by a sequence of dose-response curves for stones, irritation, hyperplasia, and tumors, each curve higher on the dose scale than its immediate precursor. A nonlinear dose-response assessment considers more than tumors as it identifies a dose where events that can lead to tumor development would not occur. Identification of a key event does not imply that it is adverse in itself, only that it is an observable step preceding tumor development. Basing a dose-response assessment on key events is intended to protect against not only the observation of adverse effects, but also earlier damage that can lead to later tumor development.

Thus, it is most desirable to estimate a dose-response curve for the key event precipitating tumor development, and use this curve to estimate the point of departure. However, lack of quantitative information on the key event may make it necessary to use tumor data instead of key event data. In this case, the analysis of the margin of exposure must contain an estimate of the dose-response curve for tumors plus have sufficient discussion of the difference (on the dose scale) between no effect levels and effect levels for key events and for tumors. A larger margin of exposure may be needed to account for possible differences between the dose-response curves for the key events and for tumors, and to assure decision-makers that cancer risk for the heterogeneous population (including sensitive subgroups) is not appreciable.

- *Slope of the observed dose-response curve*. Have we reached a dose where tumors or (preferably) the key precursor events *would not occur*? A 10-percent incidence is typically used as a point of departure because it reflects the lowest incidence that experimental studies can typically detect. This does not, however, mean that a 10-percent incidence represents a level where tumors or the key precursor events would not occur. To account for this limitation, one needs to consider the slope of the dose-response curve, which describes how sharply the incidence declines below the point of departure. If the dose-response curve at the point of departure is relatively steep, the point of departure represents a point on the dose-response curve where occurrence of the key event(s) declines rapidly with decreasing dose. On the other hand, if the dose-response curve is

relatively shallow, then the point where the effect virtually disappears may lie far below the point of departure. In short, the margin of exposure needs to be larger if the analysis is based on a response(s) that has a shallow dose-response curve compared to an analysis based on a response with a steep dose-response curve. More guidance needs to be developed to define quantitatively what constitutes a steep versus a shallow dose-response curve.

- *Human sensitivity compared with experimental animals.* How sensitive is the *human population* compared with the tested animals? For this comparison, all doses should have already been converted to equivalent human doses, using either a physiologically based toxicokinetic model, a cross-species dosimetry model, or the default cross-species scaling factor. These dose conversions reflect interspecies differences in toxicokinetics, not toxicodynamics. When information is not sufficient to quantify human sensitivity with regard to the toxicodynamics compared with the tested animals, this uncertainty needs to be taken into account in the discussion of an adequate margin of exposure. As with noncancer assessment, the default assumption is that the most sensitive humans are more sensitive than the test animals. Depending on the data available on the sensitivity of the test species to the agent and the endpoint of concern as compared to humans, the margin of exposure decision may need to incorporate more or less conservatism.

- *Nature and extent of human variability in sensitivity.* Is there information on *sensitive individuals* that would be part of a heterogeneous human population? Pertinent information would come from human studies, since animal studies, particularly those using homogeneous animal strains, do not provide information about human variability. When information is not sufficient to quantify the extent of human variability in sensitivity, this uncertainty should be reflected in the discussion of an adequate margin of exposure (also see discussion below on human exposure).

- *Human Exposure.* The evaluation of margin of exposure also takes into account the expected pattern of human exposure to an agent including the magnitude, frequency, and duration of exposure. Some modes of action involve significant duration of exposure before tumorigenicity results. For example, stimulus of cell growth through hormonal or other signal disruption or as a result of damage from toxicity is reversible if the exposure is for a short time, since homeostasis brings a return to normal levels after cessation of

exposure. Thus, for a specialized population that is occasionally and briefly exposed to an agent with such a mode of action, an adequate of margin of exposure would be smaller than for chronic exposure. As the duration of exposure or frequency of exposure increases, an adequate margin of exposure would increase accordingly.

Furthermore, if the population exposed in a particular scenario is wholly or largely composed of a subpopulation of special concern (e.g. children) for whom evidence indicates a special sensitivity to the agent's mode of action, an adequate margin of exposure would be larger than for general population exposure.

To provide input regarding scientific considerations regarding the acceptability of a margin of exposure by the risk manager, the risk assessment along with risk characterization explicitly considers all of the hazard and dose-response and human exposure factors together. This input on the margin of exposure is not solely a composite of individual adjustment factors to account for missing data or knowledge gaps as discussed above. Rather, each case calls for individual judgment, taking all of these points as a whole. It is appropriate to provide a graphical representation of the data and dose-response modeling in the observed range, also showing exposure levels of interest to the decision-maker (See figure 3-1.). In order to provide a frame of reference, by way of comparison, a straight line extrapolation may be displayed to show what risk levels would be associated with decreasing dose, if the dose-response were linear.

3.3.3.3. Linear and Nonlinear Extrapolations

Both linear and nonlinear procedures may be used in particular cases. If a mode of action analysis finds substantial support for differing modes of action for different tumor sites, an appropriate procedure is used for each. Both procedures may also be appropriate to discuss implications of complex dose-response relationships. For example, if it is apparent that an agent is both DNA reactive and is highly active as a promotor at high doses, and there are insufficient data for modeling, both linear and nonlinear default procedures may be needed to decouple and consider the contribution of both phenomena.

3.3.3.4. Use of Toxicity Equivalence Factors and Relative Potency Estimates

A toxicity equivalence factor (TEF) procedure is one used to derive quantitative dose-response estimates for agents that are members of a category or class of agents. TEFs are based on shared characteristics that can be used to order the class members by carcinogenic potency

when cancer bioassay data are inadequate for this purpose (U.S. EPA, 1991c). The ordering is by reference to the characteristics and potency of a well-studied member or members of the class. Other class members are indexed to the reference agent(s) by one or more shared characteristics to generate their TEFs. The TEFs are usually indexed at increments of a factor of 10. Very good data may permit a smaller increment to be used. Shared characteristics that may be used are, for example, receptor-binding characteristics, results of assays of biological activity related to carcinogenicity, or structure-activity relationships.

TEFs are generated and used for the limited purpose of assessment of agents or mixtures of agents in environmental media when better data are not available. When better data become available for an agent, its TEF should be replaced or revised. Criteria for constructing TEFs are given in U.S. EPA (1991b). The criteria call for data that are adequate to support summing doses of the agents in mixtures. To date, adequate data to support use of TEF's has been found in only one class of compounds (dioxins) (U.S. EPA, 1989a).

Relative potencies can be similarly derived and used for agents with carcinogenicity or other supporting data. These are conceptually similar to TEFs, but they are less firmly based in science and do not have the same level of data to support them. They are used only when there is no better alternative.

The uncertainties associated with both TEFs and relative potencies are explained whenever they are used.

3.4. RESPONSE DATA

Response data for analysis include tumor incidence data from human or animal studies as well as data on other responses as they relate to an agent's carcinogenicity, such as effects on growth control processes or cell macromolecules or other toxic effects. Tumor incidence data are ordinarily the basis of dose-response assessment, but other response data can augment such assessment or provide separate assessments of carcinogenicity or other important effects.

Data on carcinogenic processes underlying tumor effects may be used to support biologically based or case-specific models. Other options for such data exist. If confidence is high in the linkage of a precursor effect and the tumor effect, the assessment of tumor incidence may be extended to lower dose levels by linking it to the assessment of the precursor effect (Swenberg et al., 1987). Even if a quantitative link is not appropriate, the assessment for a precursor effect may provide a view of the likely shape of the dose-response curve for tumor incidence below the range of tumor observation (Cohen and Ellwein, 1990; Choy, 1993). If responses other than tumor incidence are regarded as better representations of the carcinogenicity

of the agent, they may be used in lieu of tumor responses. For example, if it is concluded that the carcinogenic effect is secondary to another toxic effect, the dose-response for the other effect will likely be more pertinent for risk assessment. As another example, if disruption of hormone activity is the key mode of action of an agent, data on hormone activity may be used in lieu of tumor incidence data.

If adequate positive human epidemiologic response data are available, they provide an advantageous basis for analysis since concerns about interspecies extrapolation do not arise. Adequacy of human exposure data for quantification is an important consideration in deciding whether epidemiologic data are the best basis for analysis in a particular case. If adequate exposure data exist in a well-designed and well-conducted epidemiologic study that detects no effects, it may be possible to obtain an upper-bound estimate of the potential human risk to provide a check on plausibility of available estimates based on animal tumor or other responses, e.g., do confidence limits on one overlap the point estimate of the other?

When animal studies are used, response data from a species that responds most like humans should be used if information to this effect exists. If this is unknown and an agent has been tested in several experiments involving different animal species, strains, and sexes at several doses and different routes of exposure, all of the data sets are considered and compared, and a judgment is made as to the data to be used to best represent the observed data and important biological features such as mode of action. Appropriate options for presenting results include:

- use of a single data set,
- combining data from different experiments (Stiteler et al., 1993; Vater et al., 1993),
- showing a range of results from more than one data set,
- showing results from analysis of more than one statistically significant tumor response based on differing modes of action,
- representing total response in a single experiment by combining animals with statistically significant tumors at more than one site, or
- a combination of these options.

The approach judged to best represent the data is presented with the rationale for the judgment, including the biological and statistical considerations involved. The following are some points to consider:

- quality of study protocol and execution,
- proportion of malignant neoplasms,
- latency of onset of neoplasia,

- number of data points to define the relationship of dose and response,
- background incidence in test animal,
- differences in range of response among species, sexes, strains,
- most sensitive responding species, and
- availability of data on related precursor events to tumor development.

Analyses of carcinogenic effects other than tumor incidence are similarly presented and evaluated for their contribution to a best judgment on how to represent the biological data for dose-response assessment.

3.5. DOSE DATA

Whether animal experiments or epidemiologic studies are the sources of data, questions need to be addressed in arriving at an appropriate measure of dose for the anticipated environmental exposure. Among these are:

- whether the dose is expressed as an environmental concentration, applied dose, or delivered dose to the target organ,
- whether the dose is expressed in terms of a parent compound, one or more metabolites, or both,
- the impact of dose patterns and timing where significant,
- conversion from animal to human doses, where animal data are used, and
- the conversion metric between routes of exposure where necessary and appropriate.

In practice, there may be little or no information on the concentration or identity of the active form at a target; being able to compare the applied and delivered doses between routes and species is the ideal, but is rarely attained. Even so, the objective is to use available data to obtain as close to a measure of internal or delivered dose as possible.

The following discussion assumes that the analyst will have data of varying detail in different cases about toxicokinetics and metabolism. Discussed below are approaches to basic data that are most frequently available, as well as approaches and judgments for improving the analysis based on additional data. The estimation of dose in human studies is tailored to the form of dose data available.

3.5.1. Interspecies Adjustment of Dose--Adult Human

When adequate data are available, the doses used in animal studies can be adjusted to equivalent human doses using toxicokinetic information on the particular agent. The methods

used should be tailored to the nature of the data on a case-by-case basis. In rare cases, it may also be possible to make adjustments based on toxicodynamic considerations. In most cases, however, there are insufficient data available to compare dose between species. In these cases, the estimate of human equivalent dose is based on science policy default assumptions. The defaults described below are modified or replaced whenever better comparative data on toxicokinetic or metabolic relationships are available. The availability and discussion of the latter also may permit reduction or discussion of uncertainty in the analysis.

For oral exposure, the default assumption is that delivered doses are related to applied dose by a power of body weight. This assumption rests on the similarities of mammalian anatomy, physiology, and biochemistry generally observed across species. This assumption is more appropriate at low applied dose concentrations where sources of nonlinearity, such as saturation or induction of enzyme activity, are less likely to occur. To derive an equivalent human oral dose from animal data, the default procedure is to scale daily applied doses experienced for a lifetime in proportion to body weight raised to the 0.75 power ($W^{0.75}$). Equating exposure concentrations in parts per million units for food or water is an alternative version of the same default procedure because daily intakes of these are in proportion to $W^{0.75}$. The rationale for this factor rests on the empirical observation that rates of physiological processes consistently tend to maintain proportionality with $W^{0.75}$. A more extensive discussion of the rationale and data supporting the Agency's adoption of this scaling factor is in U.S. EPA, 1992b. Information such as blood levels or exposure biomarkers or other data that are available for interspecies comparison are used to improve the analysis when possible.

The default procedure to derive an human equivalent concentration of inhaled particles and gases is described in U.S. EPA (1994) and Jarabek (1995a,b). The methodology estimates respiratory deposition of inhaled particles and gases and provides methods for estimating internal doses of gases with different absorption characteristics. The method is able to incorporate additional toxicokinetics and metabolism to improve the analysis if such data are available.

3.5.2. Adjustment of Dose from Adults to Children

Slope factors and unit risk estimates for lifetime exposure incorporate exposure factors that are based on adults (specifically, body weight, breathing rate, and drinking water ingestion rate). When these unit risk estimates are used to assess risks from less-than-lifetime exposure that occurs during childhood, adjustments for differences between adults and children may be

appropriate.

Inhalation unit risk estimates: Section 3.5.1 specifies that the inhalation methodology (U.S. EPA, 1994) be used for inhaled concentrations when agent-specific data are insufficient to develop a case-specific dosimetry model. The methodology incorporates exposure factors based on a 70-kg adult who breathes at a plausibly high rate of 20 m^3/d. Because children breathe more air per unit of body weight (U.S. EPA, 1998), use of adult exposure factors may not be appropriate. Consequently, inhalation unit risk estimates are adjusted to reflect a child's body weight and breathing rate. For example, the following calculation adjusts an (adult) unit risk estimate of $1x10^{-4}$ per ug/m^3 so that it applies to a 9-kg infant who breathes 4.5 m^3/d:

$(1x10^{-4}$ per ug/m^3) x (4.5 m^3/d / 20 m^3/d) / (9 kg / 70 kg) = $1.75x10^{-4}$ per ug/m^3.

For inhaled gases and aerosols, this adjustment is intended to provide the same degree of health-conservatism for children and adults. For inhaled particles, the adjustment does not take into account the different size and spacing of airways of children and adults; this difference could result in children and adults retaining particles with a different size distribution and different toxicologic properties. To reduce this uncertainty, EPA is developing a default dosimetry model for children that is based on children's inhalation parameters.

Drinking water unit risk estimates: Similarly, drinking water unit risk estimates incorporate exposure factors based on a 70-kg adult who drinks water at a plausibly high rate of 2 L/d. Because children drink more water per unit of body weight (U.S. EPA, 1997c), use of adult exposure factors may not be appropriate. Consequently, drinking water unit risk estimates will be adjusted to reflect a child's body weight and drinking water ingestion rate.

Oral slope factors: Oral slope factors incorporate a cross-species scaling factor based on equivalence of mg/$kg^{3/4}$-d (U.S. EPA, 1992b). This cross-species factor is intended to achieve equivalence in lifetime cancer risk in different mammalian species. When risks from childhood exposure are being assessed, the child's weight is not substituted for an adult weight in the cross-species scaling factor. There are several reasons why using the child's weight in the cross-species factor may not be appropriate:

- Using the child's weight instead of an adult weight assumes that children have faster metabolism, leading to faster clearance, smaller body burdens, and smaller risks. Although children generally metabolize and eliminate many chemicals faster than adults, this is not true in all cases (Renwick, 1998).
- The data supporting the 3/4-power factor pertain to cross-species equivalence, a

fundamentally different question from determining equivalence across different life stages of a single species.

- Although exposure may begin during childhood, subsequent events that complete the carcinogenesis process may continue into adulthood.

Using an adult body weight is also a science policy choice that provides some degree of health-conservatism for children in view of the uncertainties in extrapolating risks to children. Quantitatively, the effect of this choice is rather modest; for example, basing the scaling factor on a 70-kg adult instead of a 10-kg child results in risk estimates that are 1.6 times higher ($[70/10]^{[1-3/4]} = 1.6$).

Dermal exposure: The risk of distal-site cancers from the fraction of a dermal exposure that is systemically absorbed is sometimes assessed by reducing the oral slope factor by a dermal absorption factor that reflects the ratio of absorption by the dermal route to absorption by the oral route. Use of a dermal absorption factor based on adults could increase the uncertainty in a risk assessment of childhood exposure. Neonates, especially premature infants, have much greater skin absorption than older children or adults (Schilter et al., 1996).

The risk of skin cancer from dermal exposure, in particular, from the fraction that remains on the skin and is not systemically absorbed, has generally not been addressed because methods to do so have not been developed. In order to assess children's risks from this important pathway, methodological research is needed in this area.

3.5.3. Toxicokinetic Analyses

Physiologically based mathematical models are potentially the most comprehensive way to account for toxicokinetic processes affecting dose. Models build on physiological compartmental modeling and attempt to incorporate the dynamics of tissue perfusion and the kinetics of enzymes involved in metabolism of an administered compound.

A comprehensive model requires the availability of empirical data on the carcinogenic activity contributed by parent compound and metabolite or metabolites and data by which to compare kinetics of metabolism and elimination between species. A discussion of issues of confidence accompanies presentation of model results (Monro, 1992). This includes considerations of model validation and sensitivity analysis that stress the predictive performance of the model. When a delivered dose measure is used in animal to human extrapolation of dose-response data, the assessment should discuss the confidence in the assumption that the

toxicodynamics of the target tissue(s) will be the same in both species. Toxicokinetic data can improve dose-response assessment by accounting for sources of change in proportionality of applied to internal or delivered dose at various levels of applied dose. Many of the sources of potential nonlinearity involve saturation or induction of enzymatic processes at high doses. An analysis that accounts for nonlinearity (for instance, due to enzyme saturation kinetics) can assist in avoiding overestimation or underestimation of low dose-response otherwise resulting from extrapolation from a sublinear or supralinear part of the experimental dose-response curve (Gillette, 1983). Toxicokinetic processes tend to become linear at low doses, an expectation that is more robust than low-dose linearity of response (Hattis, 1990). Accounting for toxicokinetic nonlinearities allows better description of the shape of the curve at relatively high levels of dose in the range of observation, but cannot determine linearity or nonlinearity of response at low dose levels (Lutz, 1990a; Swenberg et al., 1987).

Toxicokinetic modeling results may be presented as the preferred method of estimating human equivalent dose or in parallel discussion with default assumptions depending on relative confidence in the modeling.

3.5.4. Route-to-Route Extrapolation

Judgments frequently need to be made about the carcinogenicity of an agent through a route of exposure different than the one in the underlying studies. For example, exposures of interest may be through inhalation of an agent tested primarily through animal feeding studies or through ingestion of an agent that showed positive results in human occupational studies from inhalation exposure.

Route-to-route extrapolation has both qualitative and quantitative aspects. For the qualitative aspect, the assessor weighs the degree to which positive results through one route of exposure in human or animal studies support a judgment that similar results would have been observed in appropriate studies using the route of exposure of interest. In general, confidence in making such a judgment is strengthened when the tumor effects are observed at a site distant from the portal of entry and when absorption through the route of exposure of interest is similar to absorption via the tested routes. In the absence of contrary data, the qualitative default assumption is that, if the agent is absorbed by a route to give an internal dose, it may be carcinogenic by that route. (See section 2.7.1.)

When a qualitative extrapolation can be supported, quantitative extrapolation may still be

problematic in the absence of adequate data. The differences in biological processes among routes of exposure (oral, inhalation, dermal) can be great because of, for example, first-pass effects and differing results from different exposure patterns. There is no generally applicable method for accounting for these differences in uptake processes in quantitative route-to-route extrapolation of dose-response data in the absence of good data on the agent of interest. Therefore, route-to-route extrapolation of dose data relies on a case-by-case analysis of available data. When good data on the agent itself are limited, an extrapolation analysis can be based on expectations from physical and chemical properties of the agent, properties and route-specific data on structurally analogous compounds, or in vitro or in vivo uptake data on the agent. Route-to-route uptake models may be applied if model parameters are suitable for the compound of interest. Such models are currently considered interim methods; further model development and validation is awaiting the development of more extensive data (see generally, Gerrity and Henry, 1990). For screening or hazard ranking, route-to-route extrapolation may be based on assumed quantitative comparability as a default, as long as it is reasonable to assume absorption by compared routes. When route-to-route extrapolation is used, the assessor's degree of confidence in both the qualitative and quantitative extrapolation should be discussed in the assessment and highlighted in the dose-response characterization.

3.5.5. Dose Averaging

The cumulative dose received over a lifetime, expressed as lifetime average daily dose, is generally considered an appropriate default measure of exposure to a carcinogen (Monro, 1992). The assumption is made that a high dose of a carcinogen received over a short period of time is equivalent to a corresponding low dose spread over a lifetime. While this is a reasonable default assumption based on theoretical considerations, departures from it are expected. Another approach is needed in some cases, such as when dose-rate effects are noted (e.g., formaldehyde). Cumulative dose may be replaced, as appropriate and justified by the data, with other dose measures. In such cases, modifications to the default assumption are made to take account of these effects; the rationale for the selected approach is explained.

In cases where a mode of action or other feature of the biology has been identified that has special dose implications for sensitive subpopulations (e.g., differential effects by sex or disproportionate impacts of early-life exposure), these are explained and are recorded to guide exposure assessment and risk characterization. Special problems arise when the human exposure

situation of concern suggests exposure regimens (e.g., route and dosing schedule) that are substantially different from those used in the relevant animal studies. These issues are explored and pointed out for attention in the exposure assessment and risk characterization.

3.6. DISCUSSION OF UNCERTAINTIES

The exploration of significant uncertainties in data for dose and response and in extrapolation procedures is part of the assessment. The presentation distinguishes between model uncertainty and parameter uncertainty. Model uncertainty is an uncertainty about a basic biological question. For example, a default, linear dose-response extrapolation may have been made based on tumor and other key evidence supporting the view that the model for an agent's mode of action is a DNA-reactive process. Discussion of the confidence in the extrapolation is appropriately done qualitatively or by showing results for alternatives that are equally plausible. It is not useful, for example, to conduct quantitative uncertainty analysis running multiple forms of linear models. This would obviate the function of the policy default.

Parameter uncertainties deal with numbers representing statistical or analytical measures of variance or error in data or estimates. Uncertainties in parameters are described quantitatively, if practicable, through sensitivity analysis and statistical uncertainty analysis. With the recent expansion of readily available computing capacity, computer methods are being adapted to create simulated biological data that are comparable with observed information. These simulations can be used for sensitivity analysis, for example, to analyze how small, plausible variations in the observed data could affect dose-response estimates. These simulations can also provide information about experimental uncertainty in dose-response estimates, including a distribution of estimates that are compatible with the observed data. Because these simulations are based on the observed data, they cannot assist in evaluating the extent to which the observed data as a whole are idiosyncratic rather than typical of the true situation. If quantitative analysis is not possible, significant parameter uncertainties are described qualitatively. In either case, the discussion highlights uncertainties that are specific to the agent being assessed, as distinct from those that are generic to most assessments.

Estimation of the applied dose in a human study has numerous uncertainties such as the exposure fluctuations that humans experience compared with the controlled exposures received by animals on test. In a prospective cohort study, there is opportunity to monitor exposure and human activity patterns for a period of time that supports estimation of applied dose (U.S. EPA,

1992a). In a retrospective study, exposure may be based on monitoring data but is often based on human activity patterns and levels reconstructed from historical data, contemporary data, or a combination of the two. Such reconstruction is accompanied by analysis of uncertainties considered with sensitivity analysis in the estimation of dose (Wyzga, 1988; U.S. EPA, 1986a). These uncertainties can also be assessed for any confounding factor for which a quantitative adjustment of dose-response data is made (U.S. EPA, 1984).

3.7. TECHNICAL Dose-response CHARACTERIZATION

As with hazard characterization, the dose-response characterization serves the dual purposes of presenting a technical characterization of the assessment results and supporting the risk characterization.

The characterization presents the results of analyses of dose data, of response data, and of dose-response. When alternative approaches are plausible and persuasive in selecting dose data, response data, or extrapolation procedures, the characterization follows the alternative paths of analysis and presents the results. The discussion covers the question of whether any should be preferred over others because it (or they) better represents the available data or corresponds to the view of the mechanism of action developed in the hazard assessment. The results for different tumor types by sex and species are provided along with the one(s) preferred. Similarly, results for responses other than tumor incidence are shown if appropriate.

Numerical dose-response estimates are presented to one significant figure to prevent an inappropriate sense of high precision. However, since rounding can introduce significant errors in a calculation, the rounding should be performed explicitly in the presentation of results; the actual calculations are not done with intermediate rounding. Numbers are qualified as to whether they represent central tendency or upper bounds and whether the method used is inherently more likely to overestimate or underestimate (Krewski et al., 1984).

In cases where a mode of action or other feature of the biology has been identified that has special implications for early-life exposure, differential effects by sex, or other concerns for sensitive subpopulations, these are explained. Similarly, any expectations that high dose-rate exposures may alter the risk picture for some portion of the population are described. These and other perspectives are recorded to guide exposure assessment and risk characterization. Whether the lifetime average daily dose or another measure of dose should be considered for differing exposure scenarios is discussed.

Uncertainty analyses, qualitative or quantitative if possible, are highlighted in the characterization.

The dose-response characterization routinely includes the following, as appropriate for the data available:

- identification of the kinds of data available for analysis of dose and response and for dose-response assessment,
- results of assessment as above,
- explanation of analyses in terms of quality of data available,
- selection of study/response and dose metric for assessment,
- discussion of implications of variability in human susceptibility, including for susceptible subpopulation,
- applicability of results to varying exposure scenarios--issues of route of exposure, dose rate, frequency, and duration,
- discussion of strengths and limitations (uncertainties) of the data and analyses that are quantitative as well as qualitative, and
- special issues of interpretation of data, such as:
  - selecting dose data, response data, and dose-response approach(es),
  - use of meta-analysis,
  - uncertainty and quantitative uncertainty analysis.

4. TECHNICAL EXPOSURE CHARACTERIZATION

Exposure assessment is the determination (qualitative and quantitative) of the magnitude, frequency, and duration of exposure (EPA, 1992). The following section provides a brief overview of exposure assessment principles with an emphasis on issues related to carcinogenic risk assessment. The information presented here should be used in conjunction with other guidances including: the 1992 Guidelines for Exposure Assessment, the 1995 Policy and Guidance for Risk Characterization, the 1997 Exposure Factors Handbook, the 1997 Policy for Use of Probabilistic Analysis in Risk Assessments, and the 1997 Guiding Principles for Monte Carlo Analysis. In addition, program specific guidelines for exposure assessment should be consulted.

Exposure assessment generally consists of four major steps: defining the assessment questions, selecting or developing the conceptual and mathematical models, collecting data or selecting and evaluating available data, and exposure characterization. Each of these steps is briefly described below.

Defining the Assessment Questions

In providing a clear and unambiguous statement of the purpose and scope of the exposure assessment (EPA, 1997a), consider the following.

- The management objectives of the assessment will determine whether deterministic screening level analyses are adequate or whether full probabilistic exposure characterization is needed.

- Identify and include all important sources (e.g., pesticide applications), pathways (e.g., food or water), and routes (e.g., ingestion, inhalation, and dermal) of exposure in the assessment. If a particular source, pathway, or route is omitted, a clear and transparent explanation should be provided.

- Separate analyses should be conducted for each definable subgroup within the population of interest. In particular, subgroups that are believed to be highly exposed or susceptible to a particular health effect should be studied. This includes people with

certain diseases or genetic susceptibilities, and others whose behavior or physiology may lead to higher exposure or susceptibility. Consider the following examples.

- Physiological differences between men and women (e.g., body weight and inhalation rate) may lead to important differences in exposures. See, for example, the discussion in the Exposure Factors Handbook, Appendix 1A (EPA, 1997c).

- Pregnant and lactating women may have exposures that differ from the general population (e.g., slightly higher water consumption) (EPA, 1997c). Further, exposure to pregnant women may result in exposure to the developing fetus. (NAS, 1993).

- Children consume more food per body weight than adults while consuming fewer types of foods (ILSI, 1992, NAS, 1993 and EPA, 1997c). In addition, children engage in crawling and mouthing (i.e., putting hands and objects in the mouth) behaviors which can increase their exposures.

- The elderly and disabled may have important differences in their exposures due to a more sedentary lifestyle (EPA, 1997c). In addition, the health status of this group may affect their susceptibility to the detrimental effects of exposure.

For further guidance, see the Guidelines for Exposure Assessment, § 3 (EPA, 1992).

Selecting or Developing the Conceptual and Mathematical Models

Carcinogen risk assessment models are generally based on the premise that risk is proportional to total lifetime dose. Therefore, the exposure metric used for carcinogenic risk assessment is the Lifetime Average Daily Dose (LADD). The LADD is typically used in conjunction with the Cancer Slope Factor (CSF) to calculate individual excess cancer risk. It is an estimate of the daily intake of a carcinogenic agent throughout the entire life of an individual. Depending on the objectives of the assessment, the LADD may be calculated deterministically (using point estimates for each factor to derive a point estimate of the exposure) or stochastically (using probability distributions to represent each factor and such techniques as Monte Carlo analysis to derive a distribution of the LADD) (EPA, 1997b). Stochastic analyses may help to

identify certain population segments that are highly exposed and may need to be assessed as a special subgroup. For further guidance, see the Guidelines for Exposure Assessment, § 5.3.5.2 (EPA, 1992).

When the route of exposure is inhalation or dermal contact, derivation of the LADD will often require an approach to "route-to-route extrapolation." The CSF and other measures of toxicity are typically derived from oral administered doses in animal studies. Therefore, for ingestion exposures in a human population it is not usually necessary to make adjustments to account for route specific differences in absorption and uptake. However, for inhalation and dermal exposures, such adjustments may be necessary. For further guidance, see the Guidelines for Exposure Assessment, § 2.1.4 (EPA, 1992).

As discussed elsewhere in these guidelines, there may be cases where the mode of action indicates that dose rates are important in the carcinogenic process. In these cases, short term, less-than-lifetime exposure estimates may be more appropriate for risk assessment than the LADD. Such estimates could be used to calculate the margin (MOE) that exists between exposure and the point of departure derived in the dose-response assessment.

Collecting Data or Selecting and Evaluating Available Data

After the assessment questions have been defined and the conceptual and mathematical models have been developed, it is necessary to compile and evaluate existing data or, if necessary, to collect new data. Depending on the exposure scenario under consideration, data on a wide variety of exposure factors may be needed. The U.S. EPA Exposure Factors Handbook (EPA, 1997c) contains a large compilation of exposure data with some analysis and recommendations. Some of these data are organized by age groups to assist with assessing such subgroups as children. See, for example, the Exposure Factors Handbook, Volume 1, Chapter 3 (EPA, 1997c). When using these existing data, it is important to evaluate the quality of the data and the extent to which the data are representative of the population under consideration. The U.S. EPA Guidance for Data Quality Assessment (EPA, 1996) and program specific guidances can provide further assistance for evaluating existing data.

When existing data fail to provide an adequate surrogate for the needs of a particular

assessment, it will be necessary to collect new data. Such data collection efforts should be guided by the references listed above (e.g., the Guidance for Data Quality Assessment and program specific guidance). Once again, subgroups of concern are an important consideration in any data collection effort.

Exposure Characterization

The exposure characterization is a technical characterization that presents the assessment results and supports the risk characterization. It provides a statement of the purpose, scope, and approach used in the assessment, identifying the exposure scenarios and population subgroups covered. It provides estimates of the magnitude, frequency, duration, and distribution of exposures among members of the exposed population as the data permit. It identifies and compares the contribution of different sources, pathways, and routes of exposure. In particular, a qualitative discussion of the strengths and limitations (uncertainties) of the data and models are presented.

The discussion of uncertainties is a critical component of the exposure characterization. Uncertainties can arise out of problems with the conceptual and mathematical models. Uncertainties can also arise from poor data quality and data that are not quite representative of the population or scenario of interest. Consider the following examples of uncertainties.

- National data (i.e., data collected to represent the entire U.S. population) may not be representative of exposures occurring within a regional or local population.

- Use of short term data to infer chronic, lifetime exposures must be done with caution. Using short term data to estimate long term exposures has the tendency to underestimate the number of people exposed, while overestimating the exposure levels experienced by those in the upper end (i.e., above the 90th percentile) of the exposure distribution. For further guidance, refer to the Guidelines for Exposure Assessment, § 5.3.1 (EPA, 1992).
- Children's behavior may lead to relatively high but intermittent exposures (EPA, 1998). This pattern of exposure, "one that gradually declines over the developmental period and which remains relatively constant thereafter" is not accounted for in the LADD model (ILSI, 1992). Further the physiological characteristics of children may

lead to important differences in exposure. Some of these differences can be accounted for in the LADD model. For further guidance, see the Guidelines for Exposure Assessment, § 5.3.5.2 (EPA, 1992).

Overall, the exposure characterization should provide a full description of the sources, pathways, and routes of exposure. The characterization also should include a full description of the populations assessed. In particular highly exposed or susceptible subgroups should be discussed. For further guidance on the exposure characterization, consult the 1992 Guidelines for Exposure Assessment (EPA, 1992), the 1995 Policy and Guidance for Risk Characterization (EPA, 1995b and a) and EPA's Rule Writer's Guide to Executive Order 13045 (especially Attachment C: Technical Support for Risk Assessors--Suggestions for Characterizing Risks to Children) (EPA, 1999).

5. RISK CHARACTERIZATION

5.1. PURPOSE

EPA has developed general guidance on risk characterization for use in all of its risk assessment activities. Administrator Carol Browner has issued a policy statement on risk characterization, the core of which is the following mandate:

> Each risk assessment prepared in support of decision making at EPA should include a risk characterization that follows the principles and reflects the values outlined in this policy. A risk characterization should be prepared in a manner that is clear, transparent, reasonable, and consistent with other risk characterizations of similar scope prepared across programs in the Agency. Further, discussion of risk in all EPA reports, presentations, decision packages, and other documents should be substantively consistent with the risk characterization. The nature of the risk characterization will depend upon the information available, the regulatory application of the risk information, and the resources (including time) available. In all cases, however, the assessment should identify and discuss all the major issues associated with determining the nature and extent of the risk and provide commentary on any constraints limiting fuller exposition. (U.S. EPA, 1995)

EPA is also developing a Risk Characterization Handbook (draft available as publication number EPA/600/R-99/025, dated March 1999), which provides detailed guidance to Agency staff. The discussion below does not attempt to duplicate this material but summarizes its applicability to carcinogen risk assessment.

The risk characterization process includes an integrative analysis of the major results of the risk assessment which is summarized for the risk manager in a nontechnical discussion that minimizes the use of technical terms. It is an appraisal of the science that informs the risk manager in his/her public health decisions, as do other decision-making analyses of economic, social, or technology issues. It also serves the needs of other interested readers. The summary is an information resource for preparation of risk communication information, but being somewhat technical, is not itself the usual vehicle for communication with every audience.

The integrative analysis brings together the assessments of hazard, dose response, and exposure to make risk estimates for the exposure scenarios of interest. This analysis is generally much more extensive than the Risk Characterization Summary. It may be peer-reviewed or subject to public comment along with the summary in preparation for an Agency decision. The integrative analysis may be titled differently by different EPA programs (e.g., “Staff Paper” for

criteria air pollutants), but it typically will identify exposure scenarios of interest in decision making and present risk analyses associated with them. Some of the analyses may concern scenarios in several media; others may examine, for example, only drinking water risks. The integrative analysis also may be the document that contains quantitative analyses of uncertainty.

The values supported by a risk characterization throughout the process are *transparency* in environmental decision making, *clarity* in communication, *consistency* in core assumptions and science policies from case to case, and *reasonableness*. While it is appropriate to err on the side of protection of health and the environment in the face of scientific uncertainty, common sense and reasonable application of assumptions and policies are essential to avoid unrealistic estimates of risk (U.S. EPA, 1995). Both integrative analyses and the Risk Characterization Summary present an integrated and balanced picture of the analysis of the hazard, dose response, and exposure. The risk analyst should provide summaries of the evidence and results and describe the quality of available data and the degree of confidence to be placed in the risk estimates. Important features include the constraints of available data and the state of knowledge, significant scientific issues, and significant science and science policy choices that were made when alternative interpretations of data existed (U.S. EPA, 1995). Choices made about using default assumptions or data in the assessment are explicitly discussed in the course of analysis, and if a choice is a significant issue, it is highlighted in the summary.

5.2. APPLICATION

Risk characterization is a necessary part of generating any Agency report on risk, whether the report is preliminary, to support allocation of resources toward further study, or comprehensive, to support regulatory decisions. In the former case, the detail and sophistication of the characterization are appropriately small in scale; in the latter case, appropriately extensive. Even if a document covers only parts of a risk assessment (hazard and dose-response analyses, for instance), the results of these are characterized.

Risk assessment is an iterative process that grows in depth and scope in stages from screening for priority making, to preliminary estimation, to fuller examination in support of complex regulatory decision making. Default assumptions are used at every stage because no database is ever complete, but they are predominant at screening stages and are used less as more data are gathered and incorporated at later stages. Various provisions in EPA-administered statutes require decisions based on findings that represent all stages of iteration. There are close to 30 provisions within the major statutes that require decisions based on risk, hazard, or exposure assessment. For example, Agency review of pre-manufacture notices under Section 5 of

the Toxic Substances Control Act relies on screening analyses, while requirements for industry testing under Section 4 of that act rely on preliminary analyses of risk or simply of exposure. At the other extreme, air quality criteria under the Clean Air Act rest on a rich data collection required by statute to undergo reassessment every few years. There are provisions that require ranking of hazards of numerous pollutants--by its nature a screening level of analysis--and other provisions that require a full assessment of risk. Given this range in the scope and depth of analyses, not all risk characterizations can or should be equal in coverage or depth. The risk assessor must carefully decide which issues in a particular assessment are important to present, choosing those that are noteworthy in their impact on results. For example, health effect assessments typically rely on animal data since human data are rarely available. The objective of characterization of the use of animal data is not to recount generic issues about interpreting and using animal data. Agency guidance documents cover these. Instead, the objective is to call out any significant issues that arose within the particular assessment being characterized and inform the reader about significant uncertainties that affect conclusions.

5.3. PRESENTATION OF RISK CHARACTERIZATION SUMMARY

The presentation is a nontechnical discussion of important conclusions, issues, and uncertainties that uses the hazard, dose-response, exposure, and integrative analyses for technical support. The primary technical supports within the risk assessment are the hazard characterization, dose-response characterization, and exposure characterization described in this guideline. The risk characterization is derived from these. The presentation should fulfill the aims outlined in the purpose section above.

5.4. CONTENT OF RISK CHARACTERIZATION SUMMARY

Specific guidance on hazard, dose response, and exposure characterization appears in previous sections. Overall, the risk characterization routinely includes the following, capturing the important items covered in hazard, dose response, and exposure characterization:

- primary conclusions about hazard, dose response, and exposure, including equally plausible alternatives;
- nature of key supporting information and analytic methods;
- risk estimates and their attendant uncertainties, including key uses of default assumptions when data are missing or uncertain;
- statement of the extent of extrapolation of risk estimates from observed data to

exposure levels of interest (i.e., margin of exposure) and its implications for certainty or uncertainty in quantifying risk;

- significant strengths and limitations of the data and analyses, including any major peer reviewers' issues;
- appropriate comparison with similar EPA risk analyses or common risks with which people may be familiar; and
- comparison with assessment of the same problem by another organization.

6. REFERENCES

Alison, R.H., Capen, C.C., Prentice, D.E. (1994) Neoplastic lesions of questionable significance to humans. Toxicol. Pathol. 22:179-186.

Allen, B.C., Crump, K.S., Shipp, A.M. (1988) Correlation between carcinogenic potency of chemicals in animals and humans. Risk Anal. 8:531-544.

Ames, B.N., Gold, L.S. (1990) Too many rodent carcinogens: mitogenesis increases mutagenesis. Science 249:970-971.

Anderson, E., Deisler, P.F., McCallum, D., St. Helaire, C., Spitzer, H.L., Strauss, H., Wilson, J.D., Zimmerman, R. (1993) Key Issues in Carcinogen Risk Assessment Guidelines. Society for Risk Analysis.

Ashby, J., Tennant, R.W. (1991) Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. Mutat. Res. 257:229-306.

Ashby, J., Tennant, R.W. (1994) Prediction of rodent carcinogenicity for 44 chemicals: results. Mutagenesis 9:7-15.

Ashby, J., Doerrer, N.G., Flamm, F.G., Harris, J.E., Hughes, D.H., Johannsen, F.R., Lewis, S.C., Krivanek, N.D., McCarthy, J.F., Moolenaar, R.J., Raabe, G.K., Reynolds, R.C., Smith, J.M., Stevens, J.T., Teta, M.J., Wilson, J.D. (1990) A scheme for classifying carcinogens. Regul. Toxicol. Pharmacol. 12:270-295.

Ashby, J., Brady, A., Elcombe, C.R., Elliott, B.M., Ishmael, J., Odum, J., Tugwood, D., Keltle, S., Purchase, I.F.H. (1994) Mechanistically based human hazard assessment of peroxisome proliferator-induced hepatocarcinogenesis. Hum. Exper. Toxicol. 13:1-117.

Auger, K.R., Carpenter, C.L., Cantley, L.C., Varticovski, L. (1989a) Phosphatidylinositol 3-kinase and its novel product, phosphatidylinositol 3-phosphate, are present in *Saccharomyces cerevisiae*. J. Biol. Chem. 264:20181-20184.

Auger, K.R., Sarunian, L.A., Soltoff, S.P., Libby, P., Cantley, L.C. (1989b) PDGF-dependent tyrosine phosphorylation stimulates production of novel polyphosphoinositides in intact cells. Cell 57:167-175.

Ayers, K.M., Torrey, C.E. and Reynolds, D.J. (1997) A transplacental carcinogenicity bioassay in CD-1 mice with zidovudine. Fundam. Appl. Toxicol. 38: 195-198.

Barnes, D.G., Daston, G.P, Evans, J.S., Jarabek, A.M., Kavlock, R.J., Kimmel, C.A., Park, C., Spitzer, H.L. (1995) Benchmark dose workshop: criteria for use of a benchmark dose to estimate a reference dose. Regul. Toxicol. Pharmacol. 21:296-306.

Barrett, J.C. (1992) Mechanisms of Action of Known Human Carcinogens. In: Mechanisms of Carcinogenesis in Risk Identification. IARC Sci. Pubs. No. 116, Lyon, France: International Agency for Research on Cancer; 115-134.

Barrett, J.C. (1993) Mechanisms of multistep carcinogenesis and carcinogen risk assessment. Environ. Health Perspect. 100:9-20.

Barrett, J.C. (1995) Role of mutagenesis and mitogenesis in carcinogenesis. Environ. Mutagen., in press.

Barrett, J. C., Lee, T. C. (1992) Mechanisms of arsenic-induced gene amplification. In: Gene amplification in mammalian cells: A comprehensive guide (ed. R. E. Kellems), Marcel Dekker, New York: 441-446.

Bayly, A.C., Roberts, R.A., Dive, C. (1994) Suppression of liver cell apoptosis in vitro by the nongenotoxic hepatocarcinogen and peroxisome proliferator nafenopin. J. Cell. Biol. 125:197-203.

Bellamy, C.O.C., Malcomson, R.D.G., Harrison, D.J., Wyllie, A.H. (1995) Cell death in health and disease: The biology and regulation of apoptosis. Seminars in Cancer Biology, Apoptosis in Oncogenesis and Chemotherapy 6:3-16.

Bianchi, A.B., Navone, N.M., Alda, C.M., Conti, C.J. (1991) Overlapping loss of heterozygosity by mitotic recombination on more chromosome 7F1-ter in skin carcinogenesis. Proc. Nat. Acad. Sci. 88:7590-7594.

Biggs, P.J., Warren, W., Venitt, S., Stratton, M.R. (1993) Does a genotoxic carcinogen contribute to human breast cancer? The value of mutational spectra in unraveling the etiology of cancer. Mutagenesis 8:275-283.

Birner, G., Albrecht, W., Neumann, H.G. (1990) Biomonitoring of aromatic amines. III: Hemoglobin binding and benzidine and some benzidine congeners. Arch. Toxicol. 64(2):97-102.

Blair, A., Burg, J., Foran, J., Gibb, H., Greenland, S., Morris, R., Raabe, G., Savitz, D., Teta, J., Wartenberg, D., Wong, O., Zimmerman, R. (1995) Guidelines for application of meta-analysis in environmental epidemiology. Regul. Toxicol. Pharmacol. 22:189-197.

Bois, F.Y., Krowech, G., Zeise, L. (1995) Modeling human interindividual variability in metabolism and risk: the example of 4-aminobiphenyl. Risk Anal. 15:205-213.

Bottaro, D.P., Rubin, J.S., Faletto, D.L., Chan, A.M.L., Kmieck, T.E., Vande Woude, G.F., Aaronson, S.A. (1991) Identification of the hepatocyte growth factor receptor as the c-*met* proto-oncogene product. Science 251:802-804.

Bouck, N. (1990) Tumor angiogenesis: the role of oncogenes and tumor suppressor genes. Cancer Cells 2:179-183.

Burek, J.D., Patrick, D.H., Gerson, R.J. (1988) Weight-of-biological evidence approach for assessing carcinogenicity. In: Grice, H.C., Cimina, J.L., eds. Carcinogenicity. New York: Springer Verlag; pp. 83-95.

Burin, G.J., Gibb, H.J. and Hill, R.N. 1995 Human bladder cancer: evidence for a potential irritation-induced mechanism. Fd. Chem. Toxicol. 33: 785-795.

Bus, J.S., Popp, J.A. (1987) Perspectives on the mechanism of action of the splenic toxicity of aniline and strucurally related compounds. Fundam. Chem. Toxicol. 25:619-626.

Callemen, C.J., Ehrenberg, L., Jansson, B., Osterman-Golkar, S., Segerback, D., Svensson, K, Wachtmeister, C.A. (1978) Monitoring and risk assessment by means of alkyl groups in hemoglobin in persons occupationally exposed to ethylene oxide. J. Environ. Pathol. Toxicol. 2:427-442.

Cantley, L.C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R., Soltoff, S. (1991) Oncogenes and signal transduction. Cell 64:281-302.

Caporaso, N., Hayes, R.B., Dosemeci, M., Hoover, R., Ayesh, R., Hetzel, M., Idle, J. (1989) Lung cancer risk, occupational exposure, and the debrisoquine metabolic phenotype. Cancer Res. 49:3675-3679.

Cavenee, W. K., Koufos, A., Hansen, M. F. (1986) Recessive mutant genes predisposing to human cancer. Mutat.

Res. 168:3-14.

Chang, C. C., Jone, C., Trosko, J. E., Peterson, A. R., Sevanian, A. (1988) Effect of cholesterol epoxides on the inhibition of intercellular communication and on mutation induction in Chinese hamster V79 cells. Mutat. Res. 206:471-478.

Chen, C., Farland, W. (1991) Incorporating cell proliferation in quantitative cancer risk assessment: approaches, issues, and uncertainties. In: Butterworth, B., Slaga, T., Farland, W., McClain, M., eds. Chemical Induced Cell Proliferation: Implications for Risk Assessment. New York: Wiley-Liss, pp. 481-499.

Choy, W.N. (1993) A review of the dose-response induction of DNA adducts by aflatoxin B_2 and its implications to quantitative cancer-risk assessment. Mutat. Res. 296:181-198.

Clayson, D.B. (1989) Can a mechanistic rationale be provided for non-genotoxic carcinogens identified in rodent bioassays? Mutat. Res. 221:53-67.

Clayson, D.B., Mehta, R., Iverson, F. (1994) Oxidative DNA damage--The effects of certain genotoxic and operationally non-genotoxic carcinogens. Mutat. Res. 317:25-42.

Cogliano, V.J. (1986) The U.S. EPA's methodology for adjusting the reportable quantities of potential carcinogens. Proceedings of the 7th National Conference on Management of Uncontrollable Hazardous Wastes (Superfund '86). Washington, DC: Hazardous Wastes Control Institute, 182-185.

Cohen, S.W., Ellwein, L.B. (1990) Cell proliferation in carcinogenesis. Science 249:1007-1011.

Cohen, S.M., Ellwein, L.B. (1991) Genetic errors, cell proliferation and carcinogenesis. Cancer Res. 51:6493-6505.

Cohen, S.M. and Ellwein, L.B. (1991) Genetic errors, cell proliferation, and carcinogenesis. Cancer Res. 51: 6493-6505.

Cohen, S.M., Purtilo, D.T., Ellwein, L.B. (1991) Pivotal role of increased cell proliferation in human carcinogenesis. Mod. Pathol. 4:371-375.

Cohen, S.M. (1995) Role of urinary physiology and chemistry in bladder carcinogenesis. Fd. Chem. Toxicol. 33: 715-30.

Collum, R.G., Alt, F.W. (1990) Are *myc* proteins transcription factors? Cancer Cells 2:69-73.

Connolly, R.B., Andersen, M.E. (1991) Biologically based pharmacodynamic models: tools for toxicological research and risk assessment. Ann. Rev. Pharmacol. Toxicol. 31:503-523.

Cresteil, T., (1998) Onset of xenobiotic metabolism in children: toxicological implications. Food Addit. and Contam. 15, Supplement 45-51.

Cross, M., Dexter, T. (1991) Growth factors in development, transformation, and tumorigenesis. Cell 64:271-280.

D'Souza, R.W., Francis, W.R., Bruce, R.D., Andersen, M.E. (1987) Physiologically based pharmacokinetic model for ethylene chloride and its application in risk assessment. In: Pharmacokinetics in risk assessment. Drinking Water and Health. Vol. 8. Washington, DC: National Academy Press.

deThe, H., Lavau, C., Marchio, A., Chomienne, C., Degos, L., Dejean, A. (1991) The PML-RAR∝ fusion mRNA generated by the t (15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. Cell 66:675-684.

DiGeorge, A.M. (1992) The endocrine system. In Nelson Textbook of Pediatrics. Behrman, R.E., Kliegman, R.M., Nelson, W.E., et al. eds. 14th ed. W.B. Saunders, Philadelphia PA, 1397-1472.

Dourson, M.L., Stara, J.F. (1983) Regulatory history and experimental support of uncertainty (safety) factors. Regul. Toxicol. Pharmacol. 3:224-238

Enterline, P.E., Henderson, V.L., Marsh, G.M. (1987) Exposure to arsenic. Amer. J. Epidemiol. 125:929-938.

European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC). (1991) Early indicators of non-genotoxic carcinogenesis. ECETOC Monograph No. 16. Brussels: ECETOC. Printed in Mutat. Res. 248:211-374.

Faustman, E. et al. (1997) Experimental approaches to evaluate mechanisms of developmental toxicity. In Hood, R.D., ed. Handbook of Developmental Biology. Boca Raton, LA: CRC Press. pp. 13-41.

Faustman, E.M., Allen, B.C., Kavlock, R.J., Kimmel, C.A. (1994) Dose-response assessment for developmental toxicity. Fundam. Appl. Toxicol. 23:478-486.

Fearon, E., Vogelstein, B. (1990) A genetic model for colorectal tumorigenesis. Cell 61:959-967.

Federated Association of Societies of Experimental Biology (FASEB) (1994) Evaluation of evidence for the carcinogenicity of butylated hydroxyanisole (BHA). Life Sciences Research Office, Bethesda, MD. Letter from Hamilton Brown (FDA) to John Rice (FASEB), July 28, 1994. Letter from John Rice (FASEB) to Ed Arnold (FDA), August 4, 1994.

Fidler, I.J., Radinsky, R. (1990) Genetic control of cancer metastasis. J. Natl. Cancer Inst. 82:166-168.

Fisher, D.A. , Klein, A.H. (1981) Thyroid development and disorders of thyroid function in the new born. N. Engl. J. Med. 304:702-712.

Fisher, R.A. (1950) Statistical Methods for Research Workers. Edinburgh, Scotland: Oliver and Boyd.

Flamm, W.G., Winbush, J.S. (1984) Role of mathematical models in assessment of risk and in attempts to define management strategy. Fundam. Appl. Toxicol. 4:S395-S401.

Flammang, T.J., Von Tungeln, L.S., Kadlubar, F.F. and Fu, P.P. (1997) Neonatal mouse assay for tumorigenicity: alternative to the chronic rodent bioassay. Regul. Toxicol. Pharmacol. 26: 230-240.

Flynn, G.L. (1990) Physicochemical determinants of skin absorption. In: Gerrity, T.R., Henry, C.J., eds. Principles of Route to Route Extrapolation for Risk Assessment. New York: Elsevier Science Publishing Co.; pp. 93-127.

Forsburg, S.L., Nurse, P. (1991) Identification of a G1-type cyclin pug1+ in the fission yeast *Schizosaccharomyces pombe*. Nature 351:245-248.

Garfinkel, L., Silverberg, E. (1991) Lung cancer and smoking trends in the United States over the past 25 years. Cancer 41:137-145.

Gaylor, D.W., Kodell, R.L. (1980) Linear interpolation algorithm for low-dose risk assessment of toxic substances.

J. Environ. Pathol. Toxicol. 4:305-312.

Gearhart, J.P., Herzberg, G.Z. and Jeffs, R.D. (1991) Childhood urolithiasis: experiences and advances. Pediatrics. 87: 445-450.

Gerrity, T.R., Henry, C., eds. (1990) Principles of Route to Route Extrapolation for Risk Assessment. New York: Elsevier Science Publishing Co.

Gibson, D.P., Aardema, M.J., Kerckaert, G.A., Carr, G.J., Brauninger, R.M., LeBoeuf, R.A. (1995) Detection of aneuploidy-inducing carcinogens in the Syrian hamster embryo (SHE) cell transformation assay. Mutat. Res.; Genet. Toxicol. 343:7-24.

Gillette, J.R. (1983) The use of pharmacokinetics in safety testing. In: Homburger, ed. Safety Evaluation and Regulation of Chemicals 2. 2nd Int. Conf., Cambridge, MA: Karger, Basel; pp. 125-133.

Gloeckler Ries, L.A., Hankey, B.F., Harras, A. and Devesa, S.S. 1996 Cancer incidence, mortality, and patient survival in the United States. In Schottenfeld, D. and Fraumeni, J.F., Jr., eds. Cancer epidemiology and prevention. 2nd ed. New York: Oxford University. pp. 168-191

Goldey, E.S., Kehn, L.S., Lau, C. et al. (1995) Developmental exposure to polychlorinated biphenyls (Aroclor 1254) reduces circulating thyroid hormone concentrations and causes hearing deficits in rats. Toxicaol. Appl. Pharmacol. 135:77-88

Goldman, L.R. (1998) Chemicals and childrens environment: what we don't know about risks. Environ Health Perspect 106 Suppl 3:875-880.

Goldsworthy, T.L., Hanigan, M.H., and Pitot, H.C. (1986) Models of hepatocarcinogenesis in the rat--contrasts and comparisons. CRC Crit. Rev. Toxicol. 17:61-89.

Goodman, G., Wilson, R. (1991) Predicting the carcinogenicity of chemicals in humans from rodent bioassay data. Environ. Health Perspect. 94:195-218.

Goodman, J.I., Counts, J.L. (1993) Hypomethylation of DNA: A possible nongenotoxic mechanism underlying the role of cell proliferation in carcinogenesis. Environ. Health Perspect. 101 Supp. 5:169-172.

Goodman, J.I., Ward, J.M., Popp, J.A., Klaunig, J.E., Fox, T.R. (1991) Mouse liver carcinogenesis: Mechanisms and relevance. Fundam. Appl. Toxicol. 17:651-665.

Grasso, P., Hinton, R.H. (1991) Evidence for and possible mechanisms of non-genotoxic carcinogenesis in rodent liver. Mutat. Res. 248:271-290.

Greenland, S. (1987) Quantitative methods in the review of epidemiologic literature. Epidemiol. Rev. 9:1-29.

Greenspan, F.S., Strewler, G.J. (1997) Basic and clinical endocrinology. Appleton & Lange, Stamford, CT.

Hale, G.A., Marina, N.M., Johnes-Wallace, D., Greenwarld, C.A., Jenkins, J.J., Rao, B.N., Luo, X., and Hudson, M.M. (1999) Late effects of treatment for germ cell tumors during childhood and adolescence. J. Pediatr. Hematol. Oncol. 21: 115-122.

Hammand, E.C. (1966) Smoking in relation to the death rates of one million men and women. In: Haenxzel, W., ed. Epidemiological approaches to the study of cancer and other chronic diseases. National Cancer Institute

Monograph No. 19. Washington, DC.

Hankey, B.F., Silverman, D.T. and Kaplan, R. (1993) Urinary bladder. In Miller, B.A., Ries, L.A.G., Hankey, B.F. et al., eds. SEER cancer statistics review: 1973-1990. NIH Pub. No. 93-789. Bethesda, MD: National Cancer Institute. pp. XXXVI.1-17.

Harris, C.C. (1989) Interindividual variation among humans in carcinogen metabolism, DNA adduct formation and DNA repair. Carcinogenesis 10:1563-1566.

Harris, C.C., Hollstein, M. (1993) Clinical implications of the p53 tumor suppressor gene. N. Engl. J. Med. 329:1318-1327.

Harris, H. (1990) The role of differentiation in the suppression of malignancy. J. Cell Sci. 97:5-10.

Hartwell, L.H., Kastan, M.B. (1994) Cell cycle control and cancer. Science 266:1821-1828.

Haseman, J.K. (1983) Issues: a reexamination of false-positive rates for carcinogenesis studies. Fundam. Appl. Toxicol. 3:334-339.

Haseman, J.K. (1984) Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environ. Health Perspect. 58:385-392.

Haseman, J.K. (1985) Issues in carcinogenicity testing: dose selection. Fundam. Appl. Toxicol. 5:66-78.

Haseman, J.K. (1990) Use of statistical decision rules for evaluating laboratory animal carcinogenicity studies. Fundam. Appl. Toxicol. 14:637-648.

Haseman, J.K. (1995) Data analysis: Statistical analysis and use of historical control data. Regul. Toxicol. Pharmacol. 21:52-59.

Haseman, J.K., Huff, J., Boorman, G.A. (1984) Use of historical control data in carcinogenicity studies in rodents. Toxicol. Pathol. 12:126-135.

Hatch, E.E., Palmer, J.R., Titus-Ernstoff, L., Noller, K.L., Kaufman, R.H., Mittendorf, R., Robboy, S.J., Hyer, M., Cowan, C.M., Adam, E., Colton, T., Hartge, P., Hoover, R.N. (1998) J. Am. Med. Assoc. 280:630-634.

Hatch, E.E., Palmer, J.R., Titus-Ernstoff, L., Noller., K.L. et al. (1998) Cancer risk in women exposed to diethylstilbestrol in utero. JAMA 280: 630-634.

Hattis, D. (1990) Pharmacokinetic principles for dose-rate extrapolation of carcinogenic risk from genetically active agents. Risk Anal. 10:303-316.

Havu, N., Mattsson, H., Ekman, L., Carlsson, E. (1990) Enterochromaffin-like cell carcinoids in the rat gastric mucosa following long-term administration of ranitidine. Digestion 45:189-195.

Hayward, N.K., Walker, G.J., Graham, W., Cooksley, E. (1991) Hepatocellular carcinoma mutation. Nature 352:764.

Hayward, J.J., Shane, B.S., Tindall, K.R., Cunningham, M.L. (1995) Differential in vivo mutagenicity of the carcinogen-noncarcinogen pair 2,4- and 2,6-diaminotoluene. Carcinogenesis 10:2429-2433.

Herschman, H.R. (1991) Primary response genes induced by growth factors or promoters. Ann. Rev. Biochem. 60:281-319.

Herbst, A.L., Ulfelder, H., Poskanzer, D.C. (1971) Adenocarcinoma of the vagina: association of maternal stilbestrol therapy with tumor appearance in young women. N Engl. J. Med. 284:878-881.

Hiatt, R.A., Dales, L.G., Friedman, G.D. and Hunkeler, E.M. (1982) Frequency of nephrolithiasis in a prepaid medical care program. Am. J. Epidemiol. 115:225-265

Hill, A.B. (1965) The environment and disease: Association or causation? Proc. R. Soc. Med. 58:295-300; Department of Health and Human Services (1982) The health consequences of smoking--cancer, a report of the Surgeon General. Washington. pp.17-20, 29 et seq.

Hill, R.N., Erdreich, L.S., Paynter, O.E., Roberts, P.A., Rosenthal, S.L., Wilkinson, C.F. (1989) Thyroid follicular cell carcinogenesis. Fundam. Appl. Toxicol. 12:629-697.

Hoel, D.G., Portier, C.J. (1994) Nonlinearity of dose-response functions for carcinogenicity. Environ. Health Perspect. 102 Suppl 1:109-113.

Hoel, D.G., Haseman, J.K., Hogam, M.D., Huff, J., McConnell, E.E. (1988) The impact of toxicity on carcinogenicity studies: Implications for risk assessment. Carcinogenesis 9:2045-2052.

Holliday, R. (1987) DNA methylation and epigenetic defects in carcinogenesis. Mutat. Res. 181:215-217.

Hollstein, M., Sidransky, D., Vogelstein, B., Harris, C.C. (1991) p53 mutations in human cancers. Science 253:49-53.

Hotz, K.J., Wilson, A.G.E., Thake, D.C., Roloff, M.V., Capen, C.C., Kronenberg, J.M., Brewster, D.W. (1997) Mechanism of thiazopyr-induced effects on thyroid hormone homeostasis in male Sprague-Dawley rats. Toxicol. Appl. Pharmacol. 142, 133-142.

Hsu, I.C., Metcaff, R.A., Sun, T., Welsh, J.A., Wang, N.J., Harris, C.C. (1991) Mutational hotspot in human hepatocellular carcinomas. Nature 350:427-428.

Huff, J.E. (1993) Chemicals and cancer in humans: first evidence in experimental animals. Environ. Health Perspect. 100:201-210.

Huff, J.E. (1994) Chemicals causally associated with cancers in humans and laboratory animals. A perfect concordance. In: Carcinogenesis. Waalkes, M.P., Ward, J.M., eds., New York: Raven Press; pp. 25-37.

Hulka, B.S., Margolin, B.H. (1992) Methodological issues in epidemiologic studies using biological markers. Am. J. Epidemiol. 135:122-129.

Hunter, T. (1991) Cooperation between oncogenes. Cell 64:249-270.

IARC. (1994) IARC monographs on the evaluation of carcinogenic risks to humans. Vol. 60, Some Industrial Chemicals. Lyon, France: IARC; pp. 13-33.

ICRP, (1991). Recommendations of the International Commission on Radiation Protection. Ann ICRP 1991; 21 (Publication 60): 1-20.

ILSI, International Life Sciences Institute. (1990) Similarities and Differences Between Children and Adults: Implications for Risk Assessment. Foran, J.A., Ed. ILSI Press: Washington, DC.

ILSI (1992). "Similarities & Differences Between Children & Adults; Implications for Risk Assessment." ILSI Press. Washington, DC.

ILSI, International Life Science Institute. (1995) The use of biological data in cancer risk assessment. In: Olin, S., Farland, W., Park, C., Rhomberg, L., Scheuplein, R., Starr, T., Wilson, J., eds. Low-Dose Extrapolation of Cancer Risks: Issues and Perspectives. Washington, DC: ILSI Press; pp. 45-60.

ILSI, International Life Sciences Institute. (1997) Principles for the Selection of Doses in Chronic Rodent Bioassays. Foran, J.A., Ed. ILSI Press: Washington, DC.

Ingbar, S.H., Woeber, K.A. (1981) the thyroid gland. In: Textbook of endocrinology. Williams R.H., ed. W.B. Saunders, Philadelphia PA, 117-247.

Ingram, A.J., Grasso, P. (1991) Evidence for and possible mechanisms of non-genotoxic carcinogenesis in mouse skin. Mutat. Res. 248:333-340.

International Agency for Research on Cancer (IARC). (1990) Ciclosporin. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 50. Lyon, France: IARC; pp. 77-114.

Israel, M.A. (1995) Molecular biology of childhood neoplasms. In Mendelsohn, J., Howley, P. M., Israel, M.A. and Liotta, L.A., eds. The molecular basis of cancer. Philadelphia: Saunders. pp. 294-316.

Ito, N., Shirai, T., and Hasegawa, R. (1992) Medium-term bioassays for carcinogens. In Mechanisms of Carcinogenesis in Risk Identifications (Vainio, H., Magee, PN., McGregor, D.B., and McMichael. A,J., eds.), Lyon, International Agency for Research on Cancer, pp. 353-388.

Jack, D., Poynter, D., Spurling, N.W. (1983) Beta-adrenoreceptor stimulants and mesoovarian leiomyomas in the rat. Toxicology 2:315-320.

Jarabek, A.M. (1995a) The application of dosimetry models to identify key processes and parameters for default dose-response assessment approaches. Toxicol. Lett. 79:171-184.

Jarabek, A.M. (1995b) Interspecies extrapolation based on mechanistic determinants of chemical disposition. Human Eco. Risk Assess. 1(5):641-662.

Johnson, C.M., Wilson, D.M., O'Fallon, W.M., Malek, R.S. and Kurland, L.T. (1979) Renal stone epidemiology: a 25-year study in Rochester, Minnesota. Kid. Internat. 16: 624-631.

Jones, P.A. (1986) DNA methylation and cancer. Cancer Res. 46:461-466.

Kehrer, J.P. (1993) Free radicals as mediators of tissue injury and disease. Crit. Rev. Toxicol. 23:21-48.

Kelsey, J.L., Thompson, W.D., Evans, A.S. (1986) Methods in Observational Epidemiology. New York: Oxford University Press.

Khoory, B.J., Pedrolli, A., Vecchni, S., Benini, D. and Fanos, V. (1998) Renal calculosis in pediatrics. Pediatr.

Med. Chir. 20: 367-76.

Kinzler, K.W., Nilbert, M.C., Su, L.-K., Vogelstein, B., Bryan, T.M., Levy, D.B., Smith, K.J., Preisinger, A.C., Hedge, P., McKechnie, D., Finniear, R., Markham, A., Groffen, J., Boguski, M.S., Altschul, S.J., Horii, A., Ando, H., Miyoshi, Y., Miki, Y., Nishisho, I., Nakamura, Y. (1991) Identification of FAP locus genes from chromosome 5q21. Science 253:661-665.

Kodell, R.L., Park, C.N. (1995) Linear Extrapolation in Cancer Risk Assessment. ILSI Risk Science Institute: Washington, D.C. In press.

Kraus, A.L., Munro, I.C., Orr, J.C., Binder, R.L., LeBoeuf, R.A., Williams, G.M. (1995) Benzoyl peroxide: an integrated human safety assessment for carcinogenicity. Regul. Toxicol. Pharmacol. 21:87-107.

Krewski, D., Brown, C., Murdoch, D. (1984) Determining "safe" levels of exposure: Safety factors of mathematical models. Fundam. Appl. Toxicol. 4:S383-S394.

Krewski, D., Murdoch, D.J., Withey, J.R. (1987) The application of pharmacokinetic data in carcinogenic risk assessment. In: Pharmacokinetics in Risk Assessment. Drinking water and health. Vol. 8. Washington, DC: National Academy Press, pp. 441-468.

Kripke, M.L. (1988) Immunoregulation of carcinogenesis: Past, present, and future. J. Natl. Cancer Inst. 80:722-727.

Kumar, R., Sukumar, S., Barbacid, M. (1990) Activation of *ras* oncogenes preceding the onset of neoplasia. Science 248:1101-1104.

Kushner, B.H., Heller, G., Cheung, N.K., Wollner, N., Kramer, K., Bajorin, D., Polyak, T., and Meyers, P.A. (1998) High risk of leukemia after short-term dose-intensive chemotherapy in young patients with solid tumors. J. Clin. Oncol. 16: 3016-3020.

Larsen, P.R. (1982) The thyroid. In Cecil textbook of medicine. 16th ed. Wyngaarden, J.B, Smith, L.H., eds. W.B. Saunders, Philadephia PA, 1201-1225.

Larson, R.A., LeBeau, M.M., Vardiman, J.W. and Rowley, J.D. (1996) Myeloid leukemia after hematotoxins. Environ. Health Perspect. 104: 1303-1307.

Larson, J.L., Wolf, D.C., Butterworth, B.E. (1994) Induced cytotoxicity and cell proliferation in the hepatocarcinogenicity of chloroform in female B6C3F1 mice: comparison of administration by gavage in corn oil vs. ad libitum in drinking water. Fundam. Appl. Toxicol. 22:90-102.

Levine, A.M. (1993) AIDS-related malignancies: The emerging epidemic. J. Natl. Cancer Inst. 85:1382-1397.

Levine, P.H., Hoover, R.N., eds. (1992) The emerging epidemic of non-Hodgkin's lymphoma: current knowledge regarding etiological factors. Cancer Res. 52:5426s-5574s.

Levine, A. J., Momand, J., Finlay, C. A. (1991) The p53 tumour suppressor gene. Nature 351:453-456.

Levine, A.J., Perry, M.E., Chang, A., et al. (1994) The 1993 Walter Hubert Lecture: The role of the p53 tumor-suppressor gene in tumorigenesis. Br. J. Cancer 69:409-416.

Li, J.L., Okada, S., Hamazaki, S., Ebina, Y., Midorikawa, O. (1987) Subacute nephrotoxicity and induction of

renal cell carcinoma in mice treated with ferric nitrilotriacetate. Cancer Res. 47:1867-1869.

Lijinsky, W. (1993) Species differences in carcinogenesis. In Vivo 7:65-72.

Lilienfeld, A.M., Lilienfeld, D. (1979) Foundations of Epidemiology, 2nd ed. New York: Oxford University Press.

Lindahl, T., Klein, G. Reedman, B.M., et al. (1974) Relationship between Epstein-Barr virus (EBV) DNA and the EBV-determined nuclear antigen (EBNA) in Burkitt lymphoma biopsies and other lymphoproliferative malignancies. Int. J. Cancer. 13: 764-772.

Loeb, L.A. (1991) Mutator phenotype may be required for multistage carcinogenesis. Cancer Res. 51:3075-3079.

Logothetopoulos, J.H. (1963) Growth and function of the thyroid gland in rats injected with L-thyroxine from birth to maturity. Endocrinology 73:349-352.

Lutz, W.K. (1990a) Endogenous genotoxic agents and processes as a basis of spontaneous carcinogenesis. Mutat. Res. 238:287-295.

Lutz, W.K. (1990b) Dose-response relationship and low dose extrapolation in chemical carcinogenesis. Carcinogenesis 11:1243-1247.

MacDonald, J.S., Lankas, G.R., Morrissey, R.E. (1994) Toxicokinetic and mechanistic considerations in the interpretaion of the rodent bioassay. Toxicol. Pathol. 22:124-140.

Maller, J.L. (1991) Mitotic control. Curr. Opin. Cell Biol. 3:269-275.

Maronpot, R.R., Shimkin, M.B., Witschi, H.P., Smith, L.H., and Cline, J.M. (1986) Strain A mouse pulmonary tumor test results for chemicals previously tested in National Cancer Institute carcinogenicity test. J. Natl. Cancer Inst. 76:1101-1112.

Marsman, D.S., Popp, J.A. (1994) Biological potential of basophilic hepatocellular foci and hepatic adenoma induced by the peroxisome proliferator, Wy-14,643. Carcinogenesis 15:111-117.

Mausner, J.S., Kramer, S. (1985) Epidemiology, 2nd ed. Philadelphia: W.B. Saunders Company.

McClain, R.M. (1994) Mechanistic considerations in the regulation and classification of chemical carcinogens. In: Kotsonis, F.N., Mackey, M., Hjelle, J., eds. Nutritional Toxicology. New York: Raven Press, pp. 273-304.

McConnell, E.E. (1992) Comparative response in carcinogenesis bioassay as a function of age at first exposure. In: Guzelian, P., Henry, C.J., and Olin, S.S. (eds). Similarities and difference between children and adults: implications for risk assessment. Washington, DC: ILSI Press, pp. 66-78.

McConnell, E.E., Solleveld, H.A., Swenberg, J.A., Boorman, G.A. (1986) Guidelines for combining neoplasms for evaluation of rodent carcinogenesis studies. J. Natl. Cancer Inst. 76:283-289.

McMichael, A.J. (1976) Standardized mortality ratios and the “healthy worker effect”: scratching beneath the surface. J. Occup. Med. 18:165-168.

Melnick, R.L., Huff, J.E., Barrett, J.C., Maronpot, R.R., Lucier, G., Portier, C.J. (1993a) Cell proliferation and chemical carcinogenesis: A symposium overview. Mol. Carcinog. 7:135-138.

Melnick, R.L., Huff, J.E., Barrett, J.C., Maronpot, R.R., Lucier, G., Portier, C.J. (1993b) Cell proliferation and chemical carcinogenesis. Mol. Carcinog. 7:135-138.

Meyn, M.S. (1993) High spontaneous intrachromosomal recombination rates in ataxia-telangiectasia. Science 260:1327-1330.

Modrich, P. (1994) Mismatch repair, genetic stability, and cancer. Science 266:1959-1960.

Monro, A. (1992) What is an appropriate measure of exposure when testing drugs for carcinogenicity in rodents? Toxicol. Appl. Pharmacol. 112:171-181.

Moolgavkar, S.H., Knudson, A.G. (1981) Mutation and cancer: a model for human carcinogenesis. J. Natl. Cancer Inst. 66:1037-1052.

Morrison, V., Ashby, J. (1994) A preliminary evaluation of the performance of the muta™ mouse (lacZ) and Big Blue™ (lacI) transgenic mouse mutation assays. Mutagenesis 9:367-375.

Naeve, G.S., Sharma, A., Lee, A.S. (1991) Temporal events regulating the early phases of the mammalian cell cycle. Curr. Opin. Cell Biol. 3:261-268.

National Academy of Science. (1993). "Pesticides in the Diets of Infants and Children." National Academy Press. Washington, DC.

National Research Council (NRC). (1983) Risk Assessment in the Federal Government: Managing the Process. Committee on the Institutional Means for Assessment of Risks to Public Health, Commission on Life Sciences, NRC. Washington, DC: National Academy Press.

NRC, (1990) Health effects of exposure to low levels of ionizing radiation. National Research Council. BEIR V. Washington: National Academy Press.

NRC. (1993) Pesticides in the diets of infants and children. Washington: National Academy Press.

NRC. (1993a) Pesticides in the Diets of Infants and Children. Committee on Pesticides in the Diets of Infants and Children, Commission on Life Sciences, NRC. Washington, DC: National Academy Press.

NRC. (1993b) Issues in Risk Assessment. Committee on Risk Assessment Methodology, Commission on Life Sciences, NRC. Washington, DC: National Academy Press.

NRC. (1994) Science and Judgment in Risk Assessment. Committee on Risk Assessment of Hazardous Air Pollutants, Commission on Life Sciences, NRC. Washington, DC: National Academy Press.

National Toxicology Program (NTP). (1984) Report of the Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation of the National Toxicology Program, Board of Scientific Counselors. Washington, DC: U.S. Government Printing Office. 1984-421-132:4726.

Nebreda, A.R., Martin-Zanca, D., Kaplan, D.R., Parada, L.F., Santos, E. (1991) Induction by NGF of meiotic maturation of *xenopus* oocytes expressing the *trk* proto-oncogene product. Science 252:558-561.

Newbold, R.R., Jefferson, W.N., Padilla-Burgos, E., and Bullock, B.C. (1997) Uterine carcinoma in mice treated neonatally with tamoxifen. Carcinogenesis. 18: 2293-2298.

Nicholson, J.F., Pesce, M.A. (1992) Laboratory testing and reference values (Tabel 27-2) in infants and children. In Nelson Textbook of Pediatrics. Behrman, R.E., Kliegman, R.M., Nelson, W.E., et al. eds. 14th ed. W.B. Saunders, Philadelphia PA, 1797-1826.

Nowell, P. (1976) The clonal evolution of tumor cell populations. Science 194:23-28.

Nunez, G., Hockenberry, D., McDonnell, J., Sorenson, C.M., Korsmeyer, S.J. (1991) Bcl-2 maintains B cell memory. Nature 353:71-72.

Nyandoto, P., Muhonen, T., and Joensuu, H. (1998) Second cancer among long-term survivors from Hodgkin's disease. Int. J. Radiat. Oncol. Biol. Phys. 42: 373-378.

Office of Science and Technology Policy (OSTP). (1985) Chemical carcinogens: Review of the science and its associated principles. Federal Register 50:10372-10442.

Organization for Economic Cooperation and Development (OECD). (1981) Guidelines for testing of chemicals. Carcinogenicity studies. No. 451. Paris, France.

Parkin, D.M., Stiller, C.A., Draper, G.J., et al. (1988) International incidence of childhood cancer. IARC Sci. Pub. No. 87. Lyon: International Agency for Research on Cancer.

Peltomäki, P., Aaltonen, L.A., Sisonen, P., Pylkkänen, L., Mecklin, J.-P., Järvinen, H., Green, J.S., Jass, J.R., Weber, J.L., Leach, F.S., Petersen, G.M., Hamilton, S.R., de la Chapelle, A., Vogelstein, B. (1993) Genetic mapping of a locus predisposing human colorectal cancer. Science 260:810-812.

Peto, J. (1992) Meta-analysis of epidemiological studies of carcinogenesis. In: Mechanisms of Carcinogenesis in Risk Assessment. IARC Sci. Pubs. No. 116, Lyon, France: IARC, pp. 571-577.

Peto, J., Darby, S. (1994) Radon risk reassessed. Nature 368:97-98.

Pitot, H., Dragan, Y.P. (1991) Facts and theories concerning the mechanisms of carcinogenesis. FASEB J. 5:2280-2286.

Portier, C. (1987) Statistical properties of a two-stage model of carcinogenesis. Environ. Health Perspect. 76:125-131.

Poskanzer, D. and Herbst, A. (1977) Epidemiology of vaginal adenosis and adenocarcinoma associated with exposure to stilbestrol in utero. Cancer. 39: 1892-1895.

Prahalada, S., Majka, J.A., Soper, K.A., Nett, T.M., Bagdon, W.J., Peter, C.P., Burek, J.D., MacDonald, J.S., van Zwieten, M.J. (1994) Leydig cell hyperplasia and adenomas in mice treated with finasteride, a 5α-reductase inhibitor: A possible mechanism. Fundam. Appl. Toxicol. 22:211-219.

Raff, M.C. (1992) Social controls on cell survival and cell death. Nature 356:397-400.

Rall, D.P. (1991) Carcinogens and human health: Part 2. Science 251:10-11.

Regulatory Toxicology and Pharmacology. (1996) 24:126-140

Renwick, A.G., (1993) Data-derived safety factors for the evaluation of food additives and environmental contaminants. Food Addit. And Contam. 10, 275-305.

Renwick, A.G., (1998) Toxicokinetics in infants and children in relation to the ADI and TDI. Food Addit. Contam. 15, Supplement, 17-35.

Renwick, A.G., Lazarus, N.R. (1998) Human variability and noncancer risk assessment--an analysis of the default uncertainty factor. Regul. Toxicol. Pharmacol. 27:3-20.

Rice, JM. (1979) Problems and perspective in perinatal carcinogenesis: a summary of the conference. NCI Moogr. 51:271-278.

Ries, L.A.G., Kosary, C.L., Hanky, B.F., Miller, B.A., Clegg, L., Edwards, B.K., eds. (1999) SEER cancer statistics review 1993-1998. National Cancer Institute, Bethesda, MD

Robboy, S.J. et al. (1984) Increased incidence of cervical and vaginal dysplasia in 3,980 diethylstilbestrol-exposed young women. JAMA. 252: 2979-2983.

Rothman, K.T. (1986) Modern Epidemiology. Boston: Little, Brown and Company.

Salomon, D.S., Kim, N., Saeki, T., Ciardiello, F. (1990) Transforming growth factor α--an oncodevelopmental growth factor. Cancer Cells 2:389-397.

Santos-Victoriano, M., Brouhard, B.H., and Cunningham, R.J., III. 1998. Renal Stone Disease in Children. Clin. Pediat. 37:583-599.

Schottenfeld, D., Fraumeni, J.F. Jr, (1996) Cancer epidemiology and prevention. Oxford University Press N.Y., Oxford.

Schneider, C., Gustincich, S., DelSal, G. (1991) The complexity of cell proliferation control in mammalian cells. Curr. Opin. Cell Biol. 3:276-281.

Schuller, H.M. (1991) Receptor-mediated mitogenic signals and lung cancer. Cancer Cells 3:496-503.

Schulte-Hermann, R., Bursch, W., Kraupp-Grasl, B., Oberhammer, F., Wagner, A., Jirtle, R. (1993) Cell proliferation and apoptosis in normal liver and preneoplastic foci. Environ. Health Perspect. 101 (Supp. 5):87-90.

Science Advisory Board (1997) An SAB report: Guidelines for Cancer Risk Assessment. Washington, DC, EPA-SAB-EHC-97-010.

Shelby, M.D., Zeiger, E. (1990) Activity of human carcinogens in the *Salmonella* and rodent bone-marrow cytogenetics tests. Mutat. Res. 234:257-261.

Shelby, M.D. (1994) Human germ cell mutations. Environ. Mol. Mutagen. 23 (Supp. 24):30-34.

Sidransky, D., Von Eschenbach, A., Tsai, Y.C., Jones, P., Summerhayes, I., Marshall, F., Paul, M., Green, P., Hamilton, P.F., Vogelstein, B. (1991) Identification of p53 gene mutations in bladder cancers and urine samples. Science 252:706-710.

Sidransky, D., Mikkelsen, T., Schwechheimer, K., Rosenblum, M.L., Cavanee, W., Vogelstein, B. (1992) Clonal expansion of p53 mutant cells is associated with brain tumor progression. Nature 355:846-847.

Sisk, S.C., Pluta, L.J., Bond, J.A., Recio, L. (1994) Molecular analysis of lacI mutants from bone marrow of B6C3F1 transgenic mice following inhalation exposure to 1,3-butadiene. Carcinogenesis 15(3):471-477.

Snedecor, G.W., Cochran, W.G. (1978) Statistical methods, Sixth ed. Ames, Iowa: Iowa State University Press, 593 pp.

Snodgrass, W.R. (1992) Physiological and biochemical differences between children and adults as determinants of toxic response to environmental pollutants. In Guzelian, P.S., Henry, C.J. and Olin, S.S. Similarities and differences between children and adults. Washington: ILSI Press. pp. 35-42.

Srivastava, S., Zou, Z., Pirollo, K., Blattner, W., Chang, E. (1990) Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome. Nature 348(6303):747-749.

Stapleton, F.B. (1996) Clinical approach to children with urolithiasis. Semin. Nephrol. 16: 389-97.

Stewart, B.W. (1994) Mechanisms of apoptosis: Integration of genetic, biochemical, and cellular indicators. J. Natl. Cancer Inst. 86:1286-1296.

Stiteler, W.H., Knauf, L.A., Hertzberg, R.C., Schoeny, R.S. (1993) A statistical test of compatibility of data sets to a common dose-response model. Regul. Toxicol. Pharmacol. 18:392-402.

Stitzel, K.A., McConnell, R.F., Dierckman, T.A. (1989) Effects of nitrofurantoin on the primary and secondary reproductive organs of female B6C3F1 mice. Toxicol. Pathol. 17:774-781.

Strausfeld, U., Labbe, J.C., Fesquet, D., Cavadore, J.C., Dicard, A., Sadhu, K., Russell, P., Dor'ee, M. (1991) Identification of a G1-type cyclin puc1+ in the fission yeast *Schizosaccharomyces pombe*. Nature 351:242-245.

Sukumar, S. (1989) *ras* oncogenes in chemical carcinogenesis. Curr. Top. Microbiol. Immunol. 148:93-114.

Sukumar, S. (1990) An experimental analysis of cancer: Role of *ras* oncogenes in multistep carcinogenesis. Cancer Cells 2:199-204.

Surks, M.I., Chopra, I.J., Maiash, C.N. et al. (1990) American Thyroid Association guidelines for use of laboratory tests in thyroid disorders. JAMA 263:1529-1532.

Swenberg, J.A., Richardson, F.C., Boucheron, J.A., Deal, F.H., Belinsky, S.A., Charbonneau, M., Short, B.G. (1987) High to low dose extrapolation: Critical determinants involved in the dose-response of carcinogenic substances. Environ. Health Perspect. 76:57-63.

Swierenga, S.H.H., Yamasaki, H. (1992) Performance of tests for cell transformation and gap junction intercellular communication for detecting nongenotoxic carcinogenic activity. In: Mechanisms of Carcinogenesis in Risk Identification. IARC Sci. Pubs. No. 116, Lyon, France: International Agency for Research on Cancer, pp. 165-193.

Tanaka, K., Oshimura, M., Kikiuchi, R., Seki, M., Hayashi, T, Miyaki, M. (1991) Suppression of tumorigenicity in human colon carcinoma cells by introduction of normal chromosome 5 or 18. Nature 349:340-342.

Tarone, R.E. (1982) The use of historical control information in testing for a trend in proportions. Biometrics 38:215-220.

Taylor, J.H., Watson, M.A., Devereux, T.R., Michels, R.Y., Saccomanno, G., Anderson, M. (1994) Lancet 343:86-

87.

Tennant, R.W. (1993) Stratification of rodent carcinogenicity bioassay results to reflect relative human hazard. Mutat. Res. 286:111-118.

Tennant, R.W., Ashby, J. (1991) Classification according to chemical structure, mutagenicity to *Salmonella* and level of carcinogenicity of a further 39 chemicals tested for carcinogenicity by the U.S. National Toxicology Program. Mutat. Res. 257:209-277.

Tennant, R.W., Elwell, M.R., Spalding, J.W., Griesemer, R.A. (1991) Evidence that toxic injury is not always associated with induction of chemical carcinogenesis. Mol. Carcinog. 4:420-440.

Tennant, R.W., French, J.E., Spalding, J.W. (1995) Identifying chemical carcinogens and assessing potential risk in short-term bioassays using transgenic mouse models. Environ. Health Perspect. 103:942-950.

Teta, M.J. (1999) Ethylene oxide cancer risk assessment based on epidemiological data: application of revised regulatory guidelines. Risk Anal. In press.

Thompson, T.C., Southgate, J., Kitchener, G., Land, H. (1989) Multistage carcinogenesis induced by *ras* and *myc* oncogenes in a reconstituted organ. Cell 56:917-3183.

Tinwell, H., Ashby, J. (1991) Activity of the human carcinogen MeCCNU in the mouse bone marrow micronucleus test. Environ. Molec. Mutagen. 17:152-154.

Tischler, A.S., McClain, R.M., Childers, H., Downing, J. (1991) Neurogenic signals regulate chromaffin cell proliferation and mediate the mitogenic effect of reserpine in the adult rat adrenal medulla. Lab. Invest. 65:374-376.

Tobey, R.A. (1975) Different drugs arrest cells at a number of distinct stages in G2. Nature 254:245-247.

Todd, G.C. (1986) Induction of reversibility of thyroid proliferative changes in rats given an antithyroid compound. Vet. Pathol. 23:110-117.

Tomatis, L., Aitio, A., Wilbourn, J., Shuker, L. (1989) Human carcinogens so far identified. Jpn. J. Cancer Res. 80:795-807.

Travis, C.C., McClain, T.W., Birkner, P.D. (1991) Diethylnitrosamine-induced hepatocarcinogenesis in rats: A theoretical study. Toxicol. Appl. Pharmacol. 109:289-309.

Trinchieri, A. 1996 Epidemiology of urolithiasis. Arch. Ital. Urol. Androl. 68: 203-249.

U.S. Environmental Protection Agency. (1983a) Good Laboratory Practices Standards–Toxicology Testing. Federal Register 48:53922.

U.S. Environmental Protection Agency. (1983b) Hazard Evaluations: Humans and Domestic Animals. Subdivision F. Available from: NTIS, Springfield, VA, PB 83-153916.

U.S. Environmental Protection Agency. (1983c) Health Effects Test Guidelines. Available from: NTIS, Springfield, VA, PB 83-232984.

U.S. Environmental Protection Agency. (1984) Estimation of the Public Health Risk from Exposure to Gasoline

Vapor Via the Gasoline Marketing System. Office of Health and Environmental Assessment, Washington, DC.

U.S. Environmental Protection Agency. (1986a) Health Assessment Document for Beryllium. Office of Health and Environmental Assessment, Washington, DC.

U.S. Environmental Protection Agency. (1986b) Guidelines for Carcinogen risk assessment. Federal Register 51(185):33992-34003.

U.S. Environmental Protection Agency. (1989a) Interim Procedures for Estimating Risks Associated with Exposures to Mixtures of Chlorinated Dibenzo-*p*-Dioxins and -Dibenzofurans (CDDs and CDFs) and 1989 Update. Risk Assessment Forum, Washington, DC. EPA/625/3-89/016.

U.S. Environmental Protection Agency. (1989b) Workshop on EPA Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum, Washington, DC. EPA/625/3-89/015.

U.S. Environmental Protection Agency. (1989c) Workshop on EPA Guidelines for Carcinogen Risk Assessment: Use of Human Evidence. Risk Assessment Forum, Washington, DC. EPA/625/3-90/017.

U.S. Environmental Protection Agency. (1989d) Summary of the Second Workshop Carcinogenesis Bioassay with the Dermal Route. May 18-19, 1988, Research Triangle Park, NC. EPA 560/6-89-003, available from NTIS, 5284 Port Royal Road, Springfield, VA 22161 (703-487-4650).

U.S. Environmental Protection Agency. (1991a) Pesticide Assessment Guidelines: Subdivision F, Hazard Evaluation, Human and Domestic Animals. Series 84, Mutagenicity, Addendum 9. Office of Pesticide Programs, Washington, DC. PB91-158394, 540/09-91-122.

U.S. Environmental Protection Agency. (1991b) Alpha-2u-Globulin: Association With Chemically Induced Renal Toxicity and Neoplasia in the Male Rat. Risk Assessment Forum, Washington, DC. EPA/625/3-91/019F.

U.S. Environmental Protection Agency. (1991c) Workshop Report on Toxicity Equivalency Factors for Polychlorinated Biphenyl Congeners. Risk Assessment Forum, Washington, DC. EPA/625/3-91/020.

U.S. Environmental Protection Agency. (1991d) Guidelines for Developmental Toxicity Risk Assessment. Federal Register 56(234):63798-63826.

U.S. Environmental Protection Agency. (1992a) Guidelines for Exposure Assessment. Federal Register 57(104):22888-22938.

U.S. Environmental Protection Agency. (1992b) Draft report: A Cross-Species Scaling Factor for Carcinogen Risk Assessment Based on Equivalence of $mg/kg^{3/4}$/day. Federal Register 57(109):24152-24173.

U.S. Environmental Protection Agency. (1992c) Health Assessment for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and Related Compounds (Chapters 1 through 8). Workshop Review Drafts. EPA/600/AP-92/001a through 001h.

U.S. Environmental Protection Agency. (1994a) Estimating Exposure to Dioxin-like Compounds. Office of Health and Environmental Assessment, Office of Research and Development, Washington, DC. External Review Draft, 3 vol. EPA/600/6-88/005Ca, Cb, Cc.

U.S. Environmental Protection Agency. (1994b) Report on the Workshop on Cancer Risk Assessment Guidelines Issues. Office of Research and Development, Risk Assessment Forum, Washington, DC. EPA/630/R-94/005a.

U.S. Environmental Protection Agency. (1994c) Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry. Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office, Research Triangle Park, NC. EPA/600/8-90/066F.

U.S. Environmental Protection Agency. (1995) Policy for Risk Characterization. Memorandum of Carol M. Browner, Administrator, March 21, 1995, Washington, DC.

U.S. Environmental Protection Agency. (1996a) Comparison of the effects of chemicals with combined perinatal and adult exposure vs. adult only exposure in carcinogenesis bioassays. http://www.epa.gov/pesticides/SAP/1996/index.htm

U.S. Environmental Protection Agency. (1996b) Comparison of the Effects of Chemicals With Combined Perinatal and Adult Exposure vs. Adult Only Exposure in Carcinogenesis Studies. Office of Pesticide Programs, October 1996.

U.S. Environmental Protection Agency. (1997a) A proposed OPP Policy on Determining the Need for Perinatal Carcinogenicity Testing on a Pesticide. Office of Pesticide Programs, 14 August 1997.

U.S. Environmental Protection Agency. (1997b) A Set of Scientific Issues Being Considered by the Agency in Connection With the Criteria for Requiring In-utero Cancer Studies. Office of Pesticide Programs. FIFRA Scientific Advisory Panel. September 1997 meeting report (www.epa.gov/pesticides/SAP/archive/september/finalsep.htm).

U.S. Environmental Protection Agency. (1997c) Exposure factors handbook. EPA/600/P-95/002F.

U.S. Environmental Protection Agency. (1998a) Assessment of Thyroid Follicular Cell Tumors. Office of Research and Development, Risk Assessment Forum, Washington, DC. EPA/630/R-97/002.

U.S. Environmental Protection Agency. (1998b) Peer Review Handbook. Office of Research and Development, Office of Science Policy, Washington, DC. EPA 100-B-98-001

U.S. Food and Drug Administration (1987) Sponsored Compounds in Food-producing Animals, Criteria and Procedures for Evaluating the Safety of Carcinogenic Residues. Final Rule. 21 CFR Parts 70, 500, 514 and 571.

Vahakangas, K.H., Samet, J.M., Metcalf, R.A., Welsh, J.A., Bennett, W.P., Lane, D.P., Harris, C.C. (1992) Mutation of p53 and *ras* Genes in Radon-associated Lung Cancer from Uranium Miners. Lancet 339:576-578.

Vainio, H., Magee, P., McGregor, D., McMichael, A.J. (1992) Mechanisms of Carcinogenesis in Risk Identification. IARC Sci. Pubs. No. 116. Lyon, France: IARC.

Van Sittert, N.J., De Jong, G., Clare, M.G., Davies, R., Dean, B.J., Wren, L.R., Wright, A.S. (1985) Cytogenetic, immunological, and hematological effects in workers in an ethylene oxide manufacturing plant. Br. J. Indust. Med. 42:19-26.

Vater, S.T., McGinnis, P.M., Schoeny, R.S., Velazquez, S. (1993) Biological considerations for combining carcinogenicity data for quantitative risk assessment. Regul. Toxicol. Pharmacol. 18:403-418.

Vesselinovitch, S.D., Rao, K.V.N., Mihailovich, N. (1979) Neoplastic response of mouse tissues during perinatal age periods and its significance in chemical carcingenesis. NCI Monogr 51:239.

Vogelstein, B., Fearon, E. R., Hamilton, S. R., Kern, S. E., Presinger, A. C., Leppert, M., Nakamura, Y., White,

R., Smits, A. M. M., Bos, J. L. (1988) Genetic alterations during colorectal-tumor development. N Eng J Med 319:525-532.

Weinberg, R.A. (1989) Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. Cancer Res. 49:3713-3721.

Wellstein, A., Lupu, R., Zugmaier, G., Flamm, S.L., Cheville, A.L., Bovi, P.D., Basicico, C., Lippman, M.E., Kern, F.G. (1990) Autocrine growth stimulation by secreted Kaposi fibroblast growth factor but not by endogenous basic fibroblast growth factor. Cell Growth Differ. 1:63-71.

West, G.B., Brown, J.H., Enquist, B.J. (1997) A General Model for the Origin of Allometric Scaling Laws in Biology. Science 276:122-126. Technical Comments and Responses: www. sciencemag.org/cgi/content/ full/s281/5378/751a

Woo, Y.T., Arcos, J.C. (1989) Role of structure-activity relationship analysis in evaluation of pesticides for potential carcinogenicity. In: Ragsdale, N.N., Menzer, R.E., eds. Carcinogenicity and pesticides: Principles, issues, and relationship. ACS Symposium Series No. 414. San Diego: Academic Press, pp. 175-200.

Wyzga, R.E. (1988) The role of epidemiology in risk assessments of carcinogens. Adv. Mod. Environ. Toxicol. 15:189-208.

Yamada, T., Nakamura, J., Murakami, M., Okuno, Y., Hosokawa, S., Matsuo, M., Yamada, H. (1994) The correlation of serum luteinizing hormone levels with the induction of Leydig cell tumors in rats by oxolinic acid. Toxicol. Appl. Pharmacol. 129:146-154.

Yamasaki, H. (1990) Gap junctional intercellular communication and carcinogenesis. Carcinogenesis 11:1051-1058.

Yamasaki, H. (1995) Non-genotoxic mechanisms of carcinogenesis: Studies of cell transformation and gap junctional intercellular communication. Toxicol. Lett. 77:55-61.

Zhang, K., Papageorge, A.G., Lowry, D.R. (1992) Mechanistic aspects of signaling through ras in NIH 3T3 cells. Science 257:671-674.